Topics in representation theory. IMPA 2024-0

Jethro van Ekeren

Contents

THIS IS THE FINAL VERSION AS THE CLASS ENDED.

THIS VERSION IS FROZEN FOR HISTORICAL PURPOSES.

ANOTHER (EVOLVING) BRANCH EXISTS.

1 Introduction

The course will cover some topics in the representation theory of semisimple Lie algebras. For such a Lie algebra, the Weyl character formula expresses the character of a finite dimensional irreducible module in terms of the characters of Verma modules. The irreducible module is finite dimensional exactly when it has a highest weight which is dominant integral. The extension of the character formula to more general highest weights is incredibly subtle, and was the subject of the Kazhdan-Lusztig conjecture, proved in the 1980s by Beilinson-Bernstein and Brylinski-Kashiwara. The proof revolves around the geometry of flag varieties, uses the theory of D-modules, and marks the beginning of geometric representation theory. The theorem is too difficult to prove in a one-semester course, but the goal is to learn some of the results and techniques that go into this theory.

2 Review of finite dimensional Lie algebras

Vector spaces are almost always over the field $\mathbb{C}$ of complex numbers.

Let's start at the beginning.

Definition 2.1. A Lie algebra is a vector space $\mathfrak{g}$ together with a bilinear operation $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that
$$[y, x] = -[x, y]$$
and
$$[x, [y, z]] = [[x, y], z] + [y, [x, z]].$$

Let A be an associative $\mathbb{C}$-algebra, for example $\mathrm{End}(V)$, then A becomes a Lie algebra when equipped with the bracket operation

$$[a, b] = ab - ba.$$

Let R be a commutative $\mathbb{C}$-algebra. A derivation of R is a linear map

$$D : R \to R$$

satisfying

$$D(ab) = D(a)b + aD(b).$$

For example $D = d/dx$ is a derivation of the algebra $R = \mathbb{C}[x]$. The set $Der(R)$ of derivations of R is a vector space. The composition $D_1 \circ D_2$ of derivations is not a derivation, but the bracket is

$$[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1.$$

This is another example of a Lie algebra.

Any vector space V with $[x, y] = 0$ for all $x, y \in V$ is a Lie algebra, called an abelian Lie algebra.

Definition 2.2. A Lie algebra $\mathfrak{g}$ is called simple if it contains no proper nonzero ideals. A vector subspace $\mathfrak{a} \subset \mathfrak{g}$ is a subalgebra if $[\mathfrak{a}, \mathfrak{a}] \subset \mathfrak{a}$ and an ideal if $[\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$.

Examples of finite-dimensional simple Lie algebras:

- The special linear Lie algebra

$$\mathfrak{sl}_n = \{A \in Mat_n \mid Tr(A) = 0\},$$

- The special orthogonal Lie algebra

$$\mathfrak{so}_n = \{A \in Mat_n \mid A + A^T = 0\},$$

- The symplectic Lie algebra

$$\mathfrak{sp}_{2n} = \{A \in Mat_{2n} \mid \Omega A + A^T \Omega = 0\}, \quad \text{where} \quad \Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

It turns out that these are all the finite dimensional simple Lie algebras, with 5 exceptions.

Cartan's criterion.

Definition 2.3. Let $(,) : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ be a bilinear form. We say that $(,)$ is invariant if

$$([x, y], z) = (x, [y, z])$$

for all $x, y, z \in \mathfrak{g}$.

For example if $\mathfrak{g} = sl_n$, then
$$(X, Y) = \mathrm{Tr}(XY)$$
is an invariant bilinear form. For any finite dimensional Lie algebra $\mathfrak{g}$ and $x \in \mathfrak{g}$, we write $\mathrm{ad}(x) = [x, -]$. Then the Killing form of $\mathfrak{g}$ is by definition

$$\kappa(X, Y) = \mathrm{Tr}_{\mathfrak{g}} \, \mathrm{ad}(X) \, \mathrm{ad}(Y).$$

Exercise 2.1. *Prove that an invariant bilinear form on a simple Lie algebra must in fact be symmetric.*

Definition 2.4. A semisimple Lie algebra is a direct sum of simple ones

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s.$$

Theorem 2.2 (Cartan's criterion)**.** *The finite dimensional Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is nondegenerate.*

The Killing form turns out to be key to determining the structure of simple Lie algebras.

Consider the set of diagonal matrices $\mathfrak{h} \subset \mathfrak{g} = sl_n$, and consider the bracket

$$[h, E_{ij}] = (h_i - h_j) E_{ij}.$$

This is saying that E_{ij} is an eigenvector for the $\mathrm{ad}(h)$ as $h \in \mathfrak{h}$. The eigenvalues are brought together into an object called a "weight", this is a linear function

$$\alpha_{ij} : \mathfrak{h} \to \mathbb{C}.$$

The whole of sl_n actually splits into eigenspaces for a finite number of these weights

$$sl_n = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C} E_{ij}.$$

In fact if we write ε_i for extraction of the ith diagonal entry, then $\alpha_{ij} = \varepsilon_i - \varepsilon_j$. For example for $n = 3$ we have
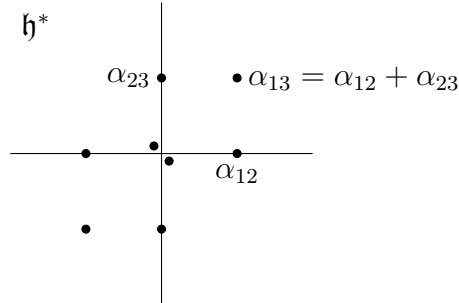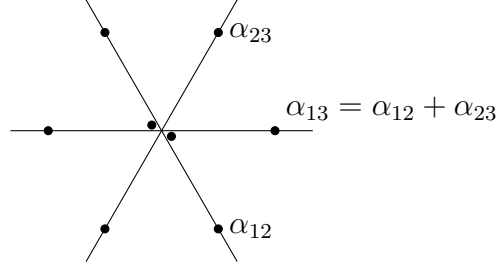

We restrict the Killing form to $(,) : \mathfrak{h} \times \mathfrak{h} \to \mathbb{C}$, and use it to define

$$\nu : \mathfrak{h} \to \mathfrak{h}^*, \quad \text{and} \quad (\alpha, \beta) = (\nu^{-1}(\alpha), \nu^{-1}(\beta)).$$

3

By the way, we use the notation $\langle \lambda, x \rangle = \langle x, \lambda \rangle = \lambda(x)$ for $x \in \mathfrak{h}$ and $\lambda \in \mathfrak{h}^*$.

If we now redraw the roots with angles dictated by the Killing form, then the figure becomes more symmetrical, as follows:

It turns out that this picture is a pretty good summary of the structure of semisimple Lie algebras in general.

Theorem 2.3. *Any finite dimensional complex (semi)simple Lie algebra $\mathfrak{g}$ has a decomposition*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathbb{C} E_\alpha,$$

where $\mathfrak{h}$ is an abelian subalgebra called the Cartan subalgebra. The dimension of $\mathfrak{h}$ is called the rank of $\mathfrak{g}$. The weights $\alpha \in \Delta \subset \mathfrak{h}^*$ are called the "roots" of $\mathfrak{g}$, i.e.,

$$[h, E_\alpha] = \alpha(h) E_\alpha \quad \text{for all } h \in \mathfrak{h}$$

and

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta},$$

(In fact $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$ whenever $\alpha + \beta$ happens to also be a root). The set of roots is invariant under operations

$$r_\alpha : \beta \mapsto \beta - 2\frac{(\alpha, \beta)}{(\alpha, \alpha)}\alpha.$$

GOT TO HERE ON 2024-01-03

The set

$$\{\beta + n\alpha \mid n \in \mathbb{Z}\} \cap (\Delta \cup \{0\})$$

is called the α-string through β. Another key result now.

Theorem 2.4. *For any pair of roots α and β the α-string through β takes the form*

$$\{\beta - p\alpha, \ldots, \beta, \ldots, \beta + q\alpha\}$$

and we have
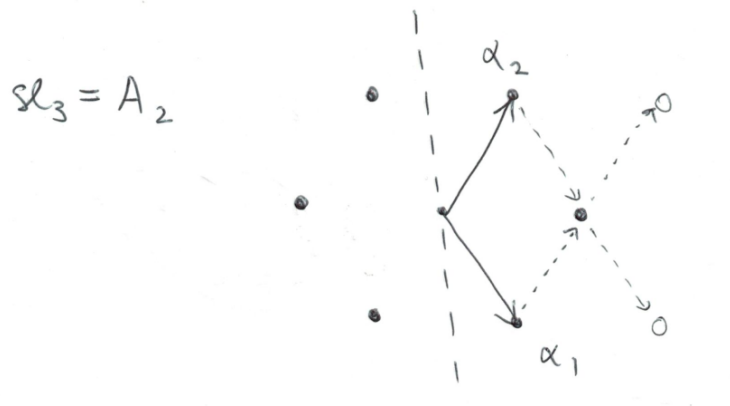
$$p - q = 2\frac{(\alpha, \beta)}{(\alpha, \alpha)}.$$

We split Δ as $\Delta_+ \cup \Delta_-$ where $\Delta_- = -\Delta_+$. Then we can define simple roots $\{\alpha_1, \ldots, \alpha_\ell\}$ as those positive roots that cannot be decomposed into a sum of other positive roots.

For simple roots α_i and α_j we have $\alpha_i - \alpha_i \notin \Delta$, therefore

$$a_{ij} = 2 \frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} = -q \leq 0$$

represents how many times one may add α_i to α_j before leaving Δ. The $\ell \times \ell$-matrix (a_{ij}) is called the Cartan matrix of $\mathfrak{g}$.
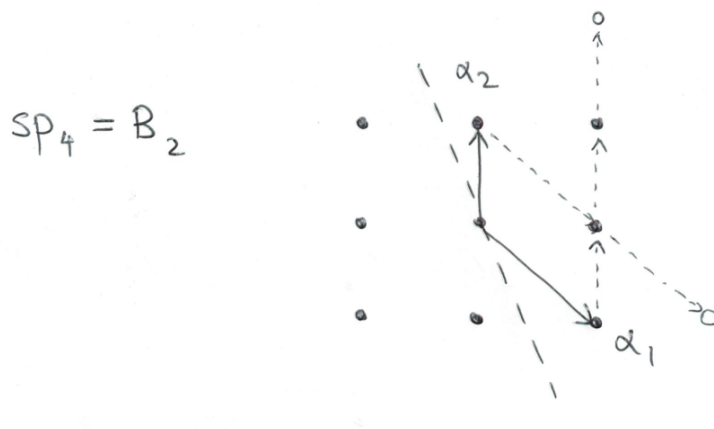
Analysing the Lie algebra sl_3 in this way:

we see that its Cartan matrix must be

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Similarly, analysing sp_4 in this way:

we see (note the importance of getting the order of the roots right) that its Cartan matrix must be

$$\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}.$$

Definition 2.5. The Weyl group of $\mathfrak{g}(A)$ is the subgroup

$$W \subset O(\mathfrak{h}^*, (,))$$

generated by the root reflections r_α.

Proposition 2.5. *In fact W is generated by the reflections $r_1, \ldots, r_\ell$ in simple roots.*

Lemma 2.6. *For each $i = 1, \ldots, \ell$ there exists some nonzero constant c such that*

$$[E_i, F_i] = c\, \nu^{-1}(\alpha_i).$$

It is standard to define

$$\alpha_i^\vee = \frac{2}{(\alpha_i, \alpha_i)} \nu^{-1}(\alpha_i),$$

and to normalise E_i and F_i so that

$$[E_i, F_i] = \alpha_i^\vee.$$

The reason for this will become clear when we review representation theory.

3 Review of modules of finite dimensional simple Lie algebras

Let $\mathfrak{g}$ be a Lie algebra. A $\mathfrak{g}$-module (also called a representation of $\mathfrak{g}$) is a vector space V and a homomorphism

$$\rho : \mathfrak{g} \to \mathrm{End}(V)$$

of Lie algebras, i.e.,

$$\rho([x, y]) = \rho(x)\rho(y) - \rho(y)\rho(x).$$

Usually we abuse notation and write xv to stand for $\rho(x)v$.

A vector subspace $W \subset V$ such that $xW \subset W$ for all $x \in \mathfrak{g}$ is called a $\mathfrak{g}$-submodule. The quotient vector space of a $\mathfrak{g}$-module by a $\mathfrak{g}$-submodule carries a natural structure of $\mathfrak{g}$-module.

PROBLEM 1: Classify all the modules of a Lie algebra $\mathfrak{g}$.

Impossible in general, so we restrict to "nice", let's say finite dimensional, modules. It's still not really possible in general. But we have

Theorem 3.1 (Weyl complete reducibility)**.** *If $\mathfrak{g}$ is finite dimensional and (semi)simple, then every finite dimensional $\mathfrak{g}$-module is isomorphic to a direct sum of irreducible $\mathfrak{g}$-modules.*

An irreducible $\mathfrak{g}$-module V is one that has no $\mathfrak{g}$-submodules other then the trivial ones (namely 0 and V itself). The proof of this theorem uses Lie algebra cohomology, we will review it soon.

PROBLEM 2: For $\mathfrak{g}$ simple, classify and describe the irreducible finite dimensional $\mathfrak{g}$-modules.

Now this is possible.

Briefly recall the answer for sl_2. Fix the basis $\{E, H, F\}$ of sl_2 as above, the relations are

$$[H, E] = 2E, \quad [H, F] = -2F, \quad [E, F] = H.$$

Theorem 3.2. *For each nonnegative integer $n \in \mathbb{Z}_+$, there is a unique irreducible module of sl_2 of dimension $n + 1$, denoted V_n. It looks as follows*

`sl2_chain_corrected.png`

GOT TO HERE ON 2024-01-04

4 Vermas, irreducibles and category $\mathcal{O}$

Let $\mathfrak{g}$ be semisimple, and let M be a $\mathfrak{g}$-module. A weight space in M is a nonzero subspace $M_\lambda \subset M$ such that

$$hv = \lambda(h)v$$

for all $h \in \mathfrak{h}$ and $v \in M$. Here $\lambda \in \mathfrak{h}^*$ is a weight. If M is a direct sum of weight spaces, we call it a weight module.

The "support" of a weight module M is the set

$$\mathrm{supp}(M) = \{\lambda \in \mathfrak{h}^* \mid M_\lambda \neq 0\}.$$

Any finite dimensional irreducible $\mathfrak{g}$-module is a weight module. Indeed if we grab a basis $\{H_1, \ldots, H_\ell\}$ of $\mathfrak{h}$ then H_1 acts in M as some endomorphism, it has an eigenspace which must be stable under H_2 since they commute, etc. So there is at least one weight space M_λ. But now one can calculate that $E_\alpha : M_\lambda \to M_{\lambda+\alpha}$ (which might be zero, but that's no problem). So M_λ is contained in a submodule M' which is a weight module. By irreducibility $M = M'$ is a weight module.

By repeatedly applying elements of $\mathfrak{g}_\alpha$ for $\alpha \in \Delta_+$ it is clear that M has one or more "highest" weights, in the obvious sense. Precisely, $\Lambda \in \mathfrak{h}^*$ is a highest weight of M if $\Lambda \in \mathrm{supp}(M)$ and for all positive roots $\alpha \in \Delta_+$ we have $\Lambda + \alpha \notin \mathrm{supp}(M)$.

Exercise 4.1. *The highest weight of an irreducible $\mathfrak{g}$-module is unique, actually.*

We recall the universal enveloping algebra $U(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$. Let $\{a_1, \ldots, a_N\}$ be a basis of $\mathfrak{g}$, then $U(\mathfrak{g})$ is a vector space spanned by words

$$a_{i_1} a_{i_2} \cdots a_{i_s}$$

with a natural multiplication given by concatenation, and unit element the empty word $1 = \emptyset$. We regard arbitrary finite strings of elements of $\mathfrak{g}$ as elements of $U(\mathfrak{g})$ by imposing distributivity of multiplication over addition. Denote by $i : \mathfrak{g} \to U(\mathfrak{g})$ the linear map sending x to the 1-letter word x.

We also impose a relation that for all $x, y \in \mathfrak{g}$ the expression $xy - yx$ in $U(\mathfrak{g})$ is to be identified with $[x, y] \in \mathfrak{g} \subset U(\mathfrak{g})$. Formally $U(\mathfrak{g})$ is the quotient of the tensor algebra $T(\mathfrak{g})$ by the 2-sided ideal generated by $x \otimes y - y \otimes x - [x, y]$ for all $x, y \in \mathfrak{g}$.

The algebra $U(\mathfrak{g})$ can be characterised by a "universal property", namely for any associative algebra A, made into a Lie algebra via $[a, b] = ab - ba$, and homomorphism φ of Lie algebras, there exists a unique homomorphism of unital associative algebras $\widetilde{\varphi}$ making the following diagram commute

$$
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{\ \ i\ \ } & U(\mathfrak{g}) \\
& \searrow{\scriptstyle\varphi} \quad \swarrow{\scriptstyle\widetilde{\varphi}} & \\
& A &
\end{array}
$$

The PBW theorem tells us that $U(\mathfrak{g})$ "looks like" a noncommutative version of the polynomial algebra $S(\mathfrak{g})$.

Theorem 4.2. *The words*

$$a_{i_1} a_{i_2} \cdots a_{i_s}$$

in which $i_1 \leq i_2 \leq \ldots \leq i_s$ *form a basis of* $U(\mathfrak{g})$.

Every $\mathfrak{g}$-module becomes a $U(\mathfrak{g})$-module in a unique way such that $i(x)v = xv$.

Definition 4.1. Let $\Lambda \in \mathfrak{h}^*$. The Verma module $M(\Lambda)$ of highest weight Λ is by definition

$$M(\Lambda) = U(\mathfrak{g}) \otimes_{U(\mathfrak{h} \oplus \mathfrak{n}_+)} \mathbb{C}_\Lambda,$$

where $\mathbb{C}_\Lambda = \mathbb{C}v_\Lambda$ is the 1-dimensional $(\mathfrak{h} \oplus \mathfrak{n}_+)$-module defined by $\mathfrak{n}_+ v_\Lambda = 0$ and $h v_\Lambda = \langle \Lambda, h \rangle v_\Lambda$ for all $h \in \mathfrak{h}$.

Recall in general if A is an associative algebra over $\mathbb{C}$ and M and N are right and left A-modules, then we have the vector space

$$M \otimes_A N = \frac{M \otimes_\mathbb{C} N}{(m \cdot a \otimes n = m \otimes a \cdot n)},$$

with variations for bimodules, etc. By the PBW theorem, as a vector space $M(\Lambda)$ is isomorphic to $U(\mathfrak{n}_-)$, and has a basis

$$f_{i_1} f_{i_2} \cdots f_{i_s} v_\Lambda, \qquad i_1 \leq \ldots \leq i_s,$$

where $f_1, f_2, \ldots f_N$ is any fixed ordered basis of $\mathfrak{n}_-$. By restriction of the action, the $\mathfrak{g}$-module $M(\Lambda)$ is also a $U(\mathfrak{n}_-)$-module. It is not difficult to see that the association

$$f_{i_1} f_{i_2} \cdots f_{i_s} \mapsto f_{i_1} f_{i_2} \cdots f_{i_s} v_\Lambda, \tag{4.1}$$

establishes an isomorphism $U(\mathfrak{n}_-) \to M(\Lambda)$ of $U(\mathfrak{n}_-)$-modules.

This isomorphism holds for all weights Λ, but as $U(\mathfrak{g})$-modules all the different $M(\Lambda)$ are distinct, as we see below, the Verma modules all have different supports, and any two isomorphic weight modules must have the same support.

Definition 4.2. Let M be a $\mathfrak{g}$-module, and let $v \in M_\lambda$ be a vector of weight λ. Then v is called a singular vector if $\mathfrak{n}_+ v = 0$

Clearly if Λ is a highest weight of M then M_Λ consists of singular vectors. But there can be other singular vectors in a $\mathfrak{g}$-module.

Like the algebra $U(\mathfrak{g})$, the module $M(\Lambda)$ can be characterised by a universal property. A singular vector v in a module V determines a unique morphism

$$M(\Lambda) \to V$$

such that $v_\Lambda \mapsto v$.

Let's introduce some notation. We denote by $Q \subset \mathfrak{h}^*$ the $\mathbb{Z}$-linear span of the set Δ of roots of $\mathfrak{g}$, i.e.,

$$Q = \{\sum_{i=1}^{\ell} k_i \alpha_i \mid k_i \in \mathbb{Z}\}$$

(We only need to use the simple roots to make linear combinations, because by definition of the notion of simple, all other roots can be written as integer linear combinations of the simple ones.) We denote by

$$Q_+ = \{\sum_{i=1}^{\ell} k_i \alpha_i \mid k_i \in \mathbb{Z}_{\geq 0}\}$$

the positive cone (equals $\mathbb{Z}_{\geq 0}$-span of Δ_+). Finally we introduce a partial order on $\mathfrak{h}^*$ by saying that

$$\lambda \geq \mu \qquad \text{if and only if } \lambda - \mu \in Q_+.$$

The support of $M(\Lambda)$ is easy to describe. It is convenient to take the $\{f_i\}$ to be a root basis of $\mathfrak{n}_-$, namely
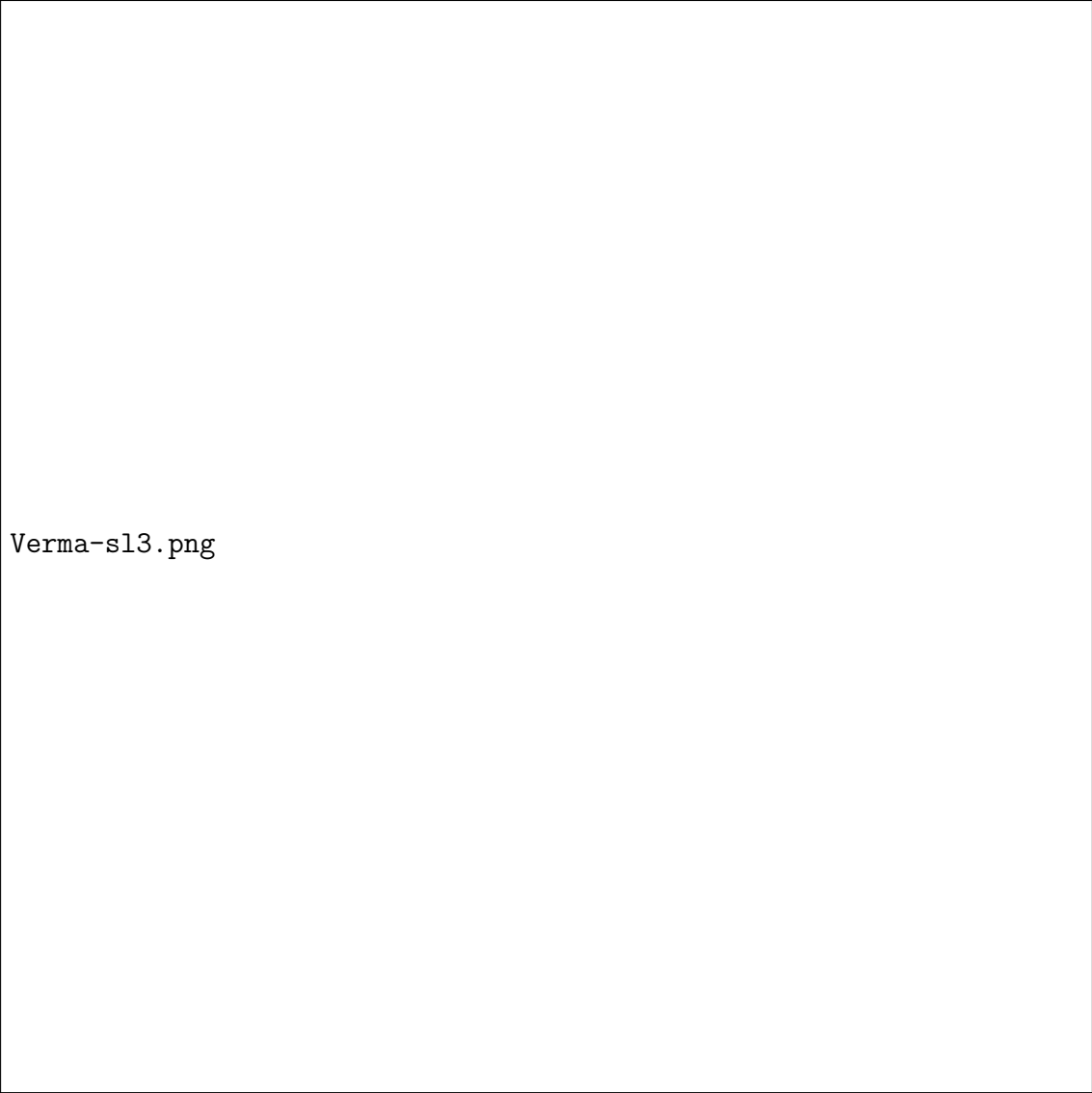
$$f_i = E_{-\alpha_i}$$

where $\{\alpha_1, \ldots, \alpha_N\}$ is Δ_+ in some chosen order. The basis vector (**??**) clearly has weight

$$\Lambda - \sum_{j=1}^{s} \alpha_{i_s}.$$

Therefore

$$\operatorname{supp} M(\Lambda) = \Lambda - Q_+.$$

It is interesting to describe the multiplicities of the weight spaces of $M(\Lambda)$ too.

With a bit of reasoning we see that

$$\dim M(\Lambda)_\Lambda = 1$$
$$\dim M(\Lambda)_{\Lambda-\alpha_1} = \dim M(\Lambda)_{\Lambda-\alpha_2} = 1$$
$$\dim M(\Lambda)_{\Lambda-\alpha_1-\alpha_2} = \dim M(\Lambda)_{\Lambda-\alpha_1-2\alpha_2} = 2.$$

Once we realise that (ultimately because of the PBW theorem) what we are counting here are basis vectors in $\mathbb{C}[f_1, f_2, f_3]$, keeping track of their weights: $\mathrm{wt}(f_1) = -\alpha_1$, $\mathrm{wt}(f_2) = -\alpha_2$, $\mathrm{wt}(f_3) = -\alpha_1 - \alpha_2$, the problem becomes easy.

We write x_1 and x_2 for the weights α_1 and α_2, so that monomials in $\mathbb{C}[f_1, f_2, f_3]$ become identified with summands in the geometric series expansion of

$$(1 - x_1^{-1})^{-1}(1 - x_2^{-1})^{-1}(1 - (x_1 x_2)^{-1})^{-1}$$

Definition 4.3. Let M be a weight $\mathfrak{g}$-module with finite dimensional weight spaces, then the formal character of M is by definition the sum

$$\mathrm{Ch}_M = \sum_{\lambda \in \mathfrak{h}^*} \dim M_\lambda e^\lambda.$$

In general the character of the Verma module $M(\Lambda)$ is easily seen to be e^Λ / R where

$$R = \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha}).$$

Example: For the finite dimensional module

$$L(3\varpi) = M(3\varpi)/J$$

of sl_2, we have write $x = e^\varpi$

$$\mathrm{Ch}_{M(3\varpi)} = x^3(1 + x^{-2} + x^{-4} + \cdots) = \frac{x^3}{1 - x^{-2}}, \qquad \mathrm{Ch}_{L(3\varpi)} = x^3 + x + x^{-1} + x^{-3}.$$

Lemma 4.3. *If $I \subseteq M(\Lambda)$ is a submodule then $I = \bigoplus_\lambda (I \cap M(\Lambda))_\lambda$.*

Proof. This was proved in my Lie algebras course 2022-2 (Lemma 14.5). $\qquad\square$

The main use of this lemma to guarantee that the sum of proper submodules of $M(\Lambda)$ is proper (each proper submodule has trivial component in weight Λ, therefore so too do sums), it follows that there exists a maximal proper submodule $J \subset M(\Lambda)$ (just take the sum of all proper submodules). We denote the quotient (which is irreducible)

$$L(\Lambda) = M(\Lambda)/J.$$

We conclude that every finite dimensional irreducible $\mathfrak{g}$-module is of this form.

Definition 4.4. The BGG Category $\mathcal{O}$ is the full subcategory of the category of all $\mathfrak{g}$-modules whose objects M satisfy the following conditions

- M is finitely generated as a $U(\mathfrak{g})$-module,

- M is a weight module, i.e.,
$$M = \bigoplus_{\lambda \in \mathfrak{h}^*} M_\lambda,$$

- the action of $\mathfrak{n}_+$ on M is locally finite, i.e., for any $v \in M$ we have
$$\dim U(\mathfrak{n}_+)v < \infty.$$

Some easy consequences of the definition. Proving these is a good exercise for those who have not seen them yet. (See [**?**, Chapter 0])

Exercise 4.4.

- *The weights of M lie in a finite union of sets of the form $\Lambda - Q_+$.*

- *The weight spaces with respect to $\mathfrak{h}$ are finite dimensional.*

GOT TO HERE ON 2024-01-08

Recall the simple roots
$$\Pi = \{\alpha_1, \ldots, \alpha_\ell\},$$
and for each one the sl_2-triple

$$E_i, \quad F_i, \quad \alpha_i^\vee = \frac{2}{(\alpha_i, \alpha_i)} \nu^{-1}(\alpha_i).$$

Now define
$$\{\varpi_1, \ldots, \varpi_\ell\} \subset \mathfrak{h}^*$$
to be the basis of $\mathfrak{h}^*$ naturally dual to

$$\Pi^\vee = \{\alpha_1^\vee, \ldots, \alpha_\ell^\vee\},$$

that is $\langle \varpi_i, \alpha_j^\vee \rangle = \delta_{ij}$. These weights are called the *fundamental weights* of $\mathfrak{g}$.

Theorem 4.5. *If $L(\Lambda)$ is finite dimensional, then $\Lambda = \sum_i k_i \varpi_i$ where all $k_i \in \mathbb{Z}_{\geq 0}$.*

By the way, we denote
$$P_+ = \{\sum_i k_i \varpi_i \mid k_i \in \mathbb{Z}_{\geq 0}\}.$$
Note that $P_+ \not\subset Q_+$ and $Q_+ \not\subset P_+$.

Proof. Let $v = v_\Lambda$ be a highest weight vector. We consider
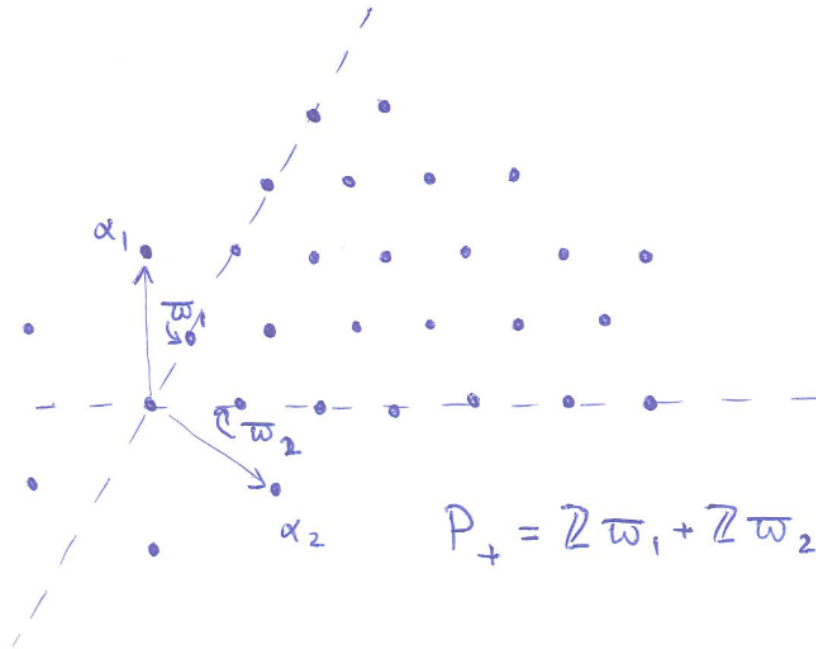
$$v, F_i v, F_i^2 v, \ldots,$$

this sequence terminates by hypothesis of finite dimensionality, which implies that

$$\alpha_i^\vee v = m v \qquad \text{for some integer } m \geq 0.$$

Therefore
$$\langle \Lambda, \alpha_i^\vee \rangle \in \mathbb{Z}_{\geq 0}.$$

Conclusion follows. $\qquad \square$

$$P_+ = \mathbb{Z}\varpi_1 + \mathbb{Z}\varpi_2$$

In fact the converse is true too.

Theorem 4.6. *If $\lambda \in P_+$ then $L(\Lambda)$ is finite dimensional.*

There should be some relation between the bases of simple roots and of fundamental weights.

Exercise 4.7. *Let $\mathfrak{g}$ be a finite dimensional simple Lie algebra, so that we know the nondegenerate invariant form $(\,,\,)$ is unique up to rescaling.*

- *Show that the coroots $\alpha_i^\vee$ are well-defined independent of the choice of $(\,,\,)$.*

- *Show that the change of basis matrix from $\{\varpi_1, \ldots, \varpi_r\}$ to $\{\alpha_1, \ldots, \alpha_r\}$ is the Cartan matrix. More precisely, that*

$$\alpha_i = \sum_{j=1}^r a_{ji}\varpi_j.$$

Let us write

$$Z(\mathfrak{g}) = \{z \in U(\mathfrak{g}) \mid zu = uz \text{ for all } u \in U(\mathfrak{g})\}$$

the centre of $U(\mathfrak{g})$. We shall now write down a very useful element of $Z(\mathfrak{g})$ called the Casimir element.

Let $(\,,\,)$ be a nondegenerate invariant bilinear form on $\mathfrak{g}$ and let $\{a_i\}$ be a basis of $\mathfrak{g}$ and $\{b_i\}$ the corresponding dual basis. That is,

$$(a_i, b_j) = \delta_{ij}.$$

The corresponding Casimir element is

$$\Omega = \sum_{i=1}^{\dim(\mathfrak{g})} a_i b_i \in U(\mathfrak{g}).$$

This element depends on the choice of form $(,)$, but having fixed such a choice it is independent of the choice of basis $\{a_i\}$. Also we recall that if $\mathfrak{g}$ is simple then any invariant bilinear form on $\mathfrak{g}$ must be a scalar multiple of the Killing form.

Proposition 4.8. *The operator Ω commutes with all $x \in \mathfrak{g}$ and thus lies in the center of $U(\mathfrak{g})$. In particular if M is an irreducible $\mathfrak{g}$-module then Ω acts as a constant in M.*

Proof. The second part is an immediate consequence of the first part and Schur's lemma. For the first part here is a sketch. Let $x \in \mathfrak{g}$ and write $[x, a_i] = \sum_j \alpha_{ij} a_j$ and similarly $[x, b_i] = \sum_j \beta_{ij} b_j$. The key is to use the dual basis relationship to show that $\beta_{ij} = -\alpha_{ij}$, and obtain a cancelation in
$$[x, \Omega] = \sum ([x, a_i] b_i + a_i [x, b_i]).$$
$\square$

Let $M = \bigoplus_\lambda M_\lambda$ be a weight module. Then since $[h, \Omega] = 0$ for all $h \in \mathfrak{h}$, we know that $\Omega(M_\lambda) \subset M_\lambda$ for all λ. If we further assume that all the weight spaces are finite dimensional (for example if $M \in \mathcal{O}$), then each weight space has a generalized eigenspace decomposition relative to Ω. So we can write

$$M = \bigoplus_{\gamma \in \mathbb{C}} M^{(\gamma)}$$

where
$$M^{(\gamma)} = \{m \in M \mid \text{there exists } N \in \mathbb{Z}_{\geq 0} \text{ such that } (\Omega - \gamma)^N m = 0\}.$$

Notice that each of these vector space summands is actually a $\mathfrak{g}$-module. So in many arguments in category $\mathcal{O}$, it is possible to assume without loss of generality that $M = M^{(\gamma)}$ for some single γ.

Let us compute the Casimir in some modules of sl_2. We take the symmetric bilinear form

$$(E, F) = 1, \quad (H, H) = 2,$$

which is invariant, and

$$\Omega = EF + FE + \frac{1}{2}H^2 = 2FE + \frac{1}{2}H(H + 2).$$

If we evaluate this operator on the highest weight vector v of weight λ (so $Ev = 0$ and $Hv = \lambda v$) then

$$\Omega = \frac{1}{2}\lambda(\lambda + 2).$$

A little more formally, we have $\alpha^\vee = H$ and $\varpi \in \mathfrak{h}^*$ characterised by $\varpi(H) = 1$. Then the highest weight $\lambda = a\varpi$ for some $a \in \mathbb{C}$, and $Hv = av$, so

$$\Omega = \frac{1}{2}a(a+2).$$

Our normalisation of the bilinear form has it that $(\alpha, \alpha) = 2$, but $\varpi = \alpha/2$, so $(\varpi, \varpi) = 1/2$. Thus really

$$\Omega = (\lambda, \lambda + 2\varpi).$$

In fact this is a special case of a general fact.

Definition 4.5. The Weyl vector of $\mathfrak{g}$ is the following sum

$$\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha.$$

Proposition 4.9. *Let $\mathfrak{g}$ be semisimple, let M be a weight $\mathfrak{g}$-module and let $v \in M_\lambda$ be a singular vector of weight $\lambda \in \mathfrak{h}^*$. Then*

$$\Omega v = (\lambda, \lambda + 2\rho)v.$$

Proof. We choose a basis $\{h_1, \ldots, h_\ell\}$ of $\mathfrak{h}$ and complete it to a basis of $\mathfrak{g}$ by adding in the root vectors E_α and F_α as α runs over Δ_+. Because of the structure of the Killing form, the dual vector to E_α is some multiple of F_α and vice versa. Let $\{h'_1, \ldots, h'_\ell\}$ be the basis of $\mathfrak{h}$ dual to $\{h_1, \ldots, h_\ell\}$.

Now we compute

$$\begin{aligned}
\Omega v &= \sum_i h_i h'_i v + \sum_{\alpha \in \Delta_+} (E_\alpha F_\alpha + F_\alpha E_\alpha)v \\
&= \sum_i \lambda(h_i)\lambda(h'_i)v + \sum_{\alpha \in \Delta_+} ([E_\alpha, F_\alpha] + 2F_\alpha E_\alpha)v \\
&= (\lambda, \lambda)v + \sum_{\alpha \in \Delta_+} \alpha^\vee v \\
&= (\lambda, \lambda)v + \sum_{\alpha \in \Delta_+} \lambda(\alpha^\vee)v
\end{aligned}$$

(I lost a factor of $2/(\alpha, \alpha)$ somehow!) This simplifies to

$$(\lambda, \lambda + 2\rho)v$$

as required. $\qquad\square$

Corollary 4.10. *Let M be a weight $\mathfrak{g}$-module, $v \in M_\lambda$ a singular vector, and $w \in U(\mathfrak{g})v$ another singular vector, this one of weight μ. Then*

$$(\mu, \mu + 2\rho) = (\lambda, \lambda + 2\rho).$$

16

Proof. We know that $\Omega v = (\lambda, \lambda + 2\rho)v$, and now because $\Omega u = u\Omega$ for all $u \in U(\mathfrak{g})$ we see that $\Omega w = (\lambda, \lambda + 2\rho)w$. But we also know that $\Omega w = (\mu, \mu + 2\rho)w$, so we are done. $\square$

Another convenient way to phrase the relation above is

$$|\mu + \rho|^2 = |\lambda + \rho|^2.$$

Proposition 4.11.

$$\rho = \sum_{i=1}^{\ell} \varpi_i.$$

Proof. Exercise. $\square$

Example of $M = M(2\varpi)$ for sl_2. There are two singular vectors: the highest weight vector in weight $\lambda = 2\varpi$, and another one in weight $\mu = -4\varpi$.

M2-M-4-inclusion.png

17

We have $\rho = \varpi$, and $\lambda + \rho = 3\varpi$ and $\mu + \rho = -3\varpi$ clearly have the same norm, as expected.

The notion of abelian category is very important. Roughly speaking an abelian category is a category $\mathcal{C}$ in which morphisms $f, g : X \to Y$ can be added and subtracted, and in which there are the notions of kernel $\ker(f) \in \mathcal{C}$ and cokernel $\mathrm{coker}(f) \in \mathcal{C}$ of a morphism $f : X \to Y$, characterised by universal properties. Notions of injective and surjective morphism (more formally known as "monomorphism" and "epimorphism", respectively) can then be recovered, as well as the "image" of a morphism, etc.

Definition 4.6. An object P in an abelian category is *projective* if for every surjection π and morphism φ as in this diagram, there exists $\widetilde{\varphi}$ completing it (i.e., such that $\pi \circ \widetilde{\varphi} = \varphi$)

$$
\begin{array}{ccc}
& & M \\
& \stackrel{\widetilde{\varphi}}{\nearrow} & \downarrow \pi \\
P & \xrightarrow{\varphi} & N
\end{array}
$$

Recall that a basic example of an abelian category is the category of all left modules over a fixed ring A. For example the category of all $\mathfrak{g}$-modules is equivalent to the category of all left $U(\mathfrak{g})$-modules, which is an abelian category.

In such categories, we can consider the free modules $A^{\oplus n}$, and it is easy to show that these are projective. In fact in such categories the projective modules are precisely the modules that occur as a summand of some free A-module.

Now it is a basic proposition that $\mathcal{O}$ is also an abelian category. But it is a proper subcategory of $U(\mathfrak{g})$-mod and in particular it is clear that it does not contain any of the free $U(\mathfrak{g})$-modules. So a priori it is not clear whether $\mathcal{O}$ contains any projective objects. But in fact it does.

Proposition 4.12. *Suppose $\Lambda \in P_+$, then the Verma module $M(\Lambda)$ is projective.*

Proof. We consider this situation

$$
\begin{array}{ccc}
& & M \\
& \stackrel{\widetilde{\varphi}}{\nearrow} & \downarrow \pi \\
M(\Lambda) & \xrightarrow{\varphi} & N
\end{array}
$$

Let $\overline{v} = \varphi(v_\Lambda) \in N$ and choose a preimage $v \in M$ of $\overline{v}$, which we can because π is surjective.

Since $\Omega v_\Lambda = (\Lambda, \Lambda + 2\rho) v_\Lambda$ we also have $\Omega \overline{v} = (\Lambda, \Lambda + 2\rho) \overline{v}$. Let us write $g = (\Lambda, \Lambda + 2\rho)$, so that $\overline{v} \in N^{(g)}$. There is also a decomposition $M = \bigoplus_{\gamma \in \mathbb{C}} M^{(\gamma)}$ and it is possible to choose $v \in M^{(g)}$.

Now consider $U(\mathfrak{n}_+)v \subset M$. This vector subspace is finite dimensional and has a weight space decomposition, so it has at least one maximal weight Λ', in other words there is some nonzero $w \in M_{\Lambda'}$ such that $w = uv$ for some $u \in U(\mathfrak{n}_+)$ and $\mathfrak{n}_+ w = 0$. Clearly $w \in M^{(g)}$, and it follows that

$$
(\Lambda, \Lambda + 2\rho) = (\Lambda', \Lambda' + 2\rho).
$$

18

We now suppose $\Lambda' \neq \Lambda$ with the aim of deriving a contradiction. In this case we have $\Lambda' = \Lambda + \beta$ for some $\beta \in Q_+$. We have

$$
\begin{aligned}
(\Lambda', \Lambda' + 2\rho) - (\Lambda, \Lambda + 2\rho) &= |\Lambda' + \rho|^2 - |\Lambda + \rho|^2 \\
&= |\Lambda + \rho + \beta|^2 - |\Lambda + \rho|^2 \\
&= 2(\Lambda + \rho, \beta) + |\beta|^2.
\end{aligned}
$$

Since $\rho = \sum_i \varpi_i$ and $\Lambda \in P_+$ we have also $\Lambda + \rho \in P_+$, and since $\beta \in Q_+$ we have $(\Lambda + \rho, \beta) \geq 0$. On the other hand $|\beta|^2 \geq 0$ with equality only for $\beta = 0$, so we have our contradiction.

Knowing now that $\Lambda' = \Lambda$, we see that $w = v$ up to a scalar, and therefore $\mathfrak{n}_+ v = 0$ or rather, v is a singular vector. By the universal property of $M(\Lambda)$ there exists a morphism $\widetilde{\varphi} : M(\Lambda) \to M$ sending $v_\Lambda \mapsto v$. This is map we wanted to construct. So $M(\Lambda)$ is projective. $\qquad \square$

We recall that for any $\mathfrak{g}$-module M, the dual vector space M^* acquires a natural $\mathfrak{g}$-module structure, via

$$
[x\varphi](m) = -\varphi(xm).
$$

One problem with (or feature of) the category $\mathcal{O}$ is that it is not closed under this operation. Firstly the dual of a vector space of infinite but countable dimension, is of uncountable dimension.

It is possible to "correct" the definition of the dual. Let

$$
\tau : \mathfrak{g} \to \mathfrak{g}
$$

be an anti-automorphism of order 2, satisfying $\tau(h) = h$ for all $h \in \mathfrak{h}$ and $\tau(\mathfrak{g}_\alpha) = \mathfrak{g}_{-\alpha}$ for all roots α. For sl_n for example, taking $\tau(X) = X^T$ would work. It is possible to construct such τ in general.

Definition 4.7. Let $M = \bigoplus_{\lambda \in \mathfrak{h}^*} M_\lambda$ be a $\mathfrak{g}$-module in category $\mathcal{O}$. The *contragredient* module $M^\vee$ is by definition $M^\vee = \bigoplus_{\lambda \in \mathfrak{h}^*} M_\lambda^*$ with the action of $\mathfrak{g}$ defined by

$$
[x\varphi](m) = \varphi(\tau(x)m).
$$

Since τ acts as the identity on $\mathfrak{h}$, we have

$$
(M^\vee)_\lambda = (M_\lambda)^* \subset M^\vee.
$$

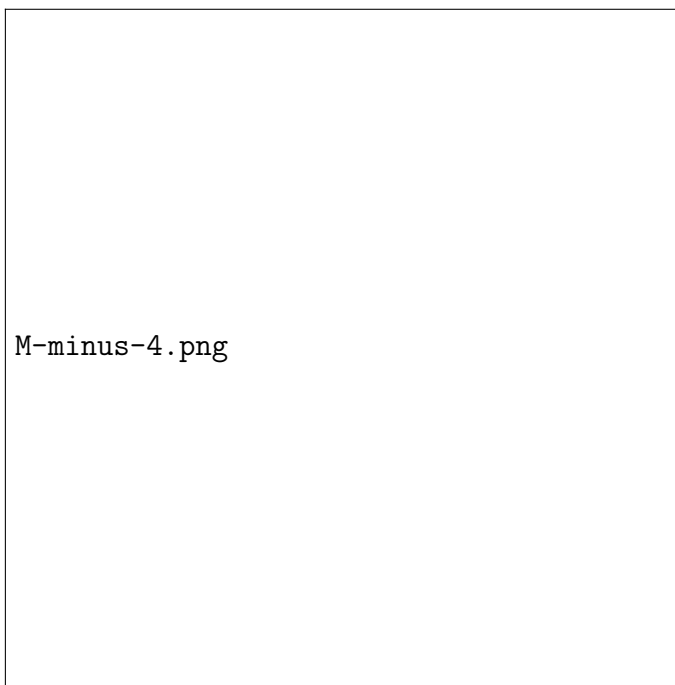We could, if we had wished to, define $M^\vee$ with $\tau = -1$, but then we would have
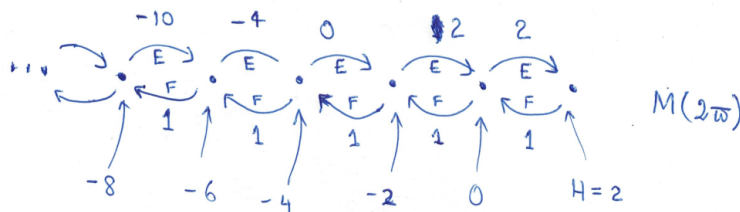
$$
(M^\vee)_\lambda = (M_{-\lambda})^* \subset M^\vee,
$$

and modules in $\mathcal{O}$ would not in general go to other modules in $\mathcal{O}$. Notice also that the characters of M and $M^\vee$ are equal.

Proposition 4.13. *If $M \in \mathcal{O}$ then also $M^\vee \in \mathcal{O}$.*

Let's do a couple of examples of contragredient modules for $\mathfrak{g} = sl_2$. Here are the Verma modules $M(2\varpi)$ and $M(-4\varpi)$ (there clearly exists an embedding $M(-4\varpi) \subset M(2\varpi)$).

M-minus-4.png

Upon taking the contragredient, we know that the set of weight spaces, and their multiplicities, remain the same. But now the coefficients of F become the coefficients of E and vice versa.
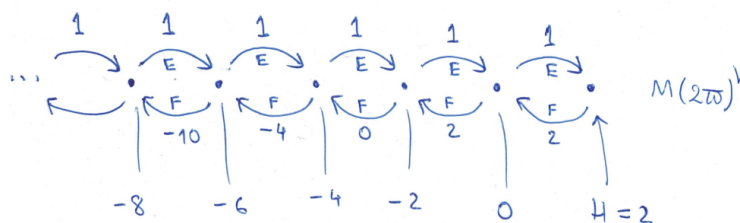
M-minus-4-dual.png

If you think about it, you will see that $M(-4\varpi)^\vee \cong M(-4\varpi)$. If we write $v_{-4}, v_{-6}, \ldots$ for the basis vectors of $M(-4\varpi)$ and $\phi_{-4}, \phi_{-6}, \ldots$ for the corresponding dual vectors which form a basis of $M(-4\varpi)^\vee$, then we can build an isomorphism inductively by $v_{-4} \mapsto \phi_{-4}$, then $v_{-6} \mapsto (-4)\phi_{-6}$, then $v_{-8} \mapsto (-40)\phi_{-8}$, and so on.

Notice that this argument does not work for $M(2\varpi)^\vee$ because we would have to map one of the basis vectors to 0, so we would not obtain an isomorphism. And indeed $M(2\varpi)^\vee \not\cong M(2\varpi)$. For instance $M(2\varpi)^\vee$ contains a copy of $L(2\varpi)$ while $M(2\varpi)$ does not.

Yet $M(2\varpi)^\vee$ is still an object of $\mathcal{O}$. Indeed the module is generated by ϕ_2 and ϕ_{-4} (actually we only need ϕ_{-4}), it is a weight module (by construction), and the action of $\mathfrak{n}_+$ is locally finite.

Got up to here on 2024-01-10

Let $\mathcal{C}$ be a $\mathbb{C}$-linear abelian category (I just mean every space $\mathrm{Hom}(X, Y)$ is a vector space over $\mathbb{C}$, and $\mathcal{C}$ is an abelian category) and $X \in \mathcal{C}$ an object. We can consider the functor

$$\mathrm{Hom}(X, -) : \mathcal{C} \to \mathbf{Vec}.$$

It is clear that this is a functor, as a morphism $f : A \to B$ induces a morphism $\mathrm{Hom}(X, A) \to \mathrm{Hom}(X, B)$ by composition $\phi \mapsto f \circ \phi$. Suppose $A \to B$ then it is basically clear that

$$\mathrm{Hom}(X, A) \subset \mathrm{Hom}(X, B),$$

as two different morphisms to A continue to be different after putting A inside B. In other words $\mathrm{Hom}(X, -)$ sends injections to injections.

The same is not true for surjections! Indeed

Definition 4.8. The object P is said to be projective if the functor $\mathrm{Hom}(P, -)$ sends surjections to surjections. That is, if for every $\varphi \in \mathrm{Hom}(P, C)$ there exists $\widetilde{\varphi} \in \mathrm{Hom}(P, B)$ such that $p \circ \widetilde{\varphi} = \varphi$

$$
\begin{array}{ccc}
B & \xrightarrow{\ p\ } & C \\
& \nwarrow \quad \diagdown & \uparrow {\scriptstyle \varphi} \\
& {\scriptstyle \widetilde{\varphi}} \diagdown & \\
& & P
\end{array}
$$

This is usually phrased in the language of short exact sequences (SES). A SES is $A \subset B$ and $C \cong B/A$, or equivalently $p : B \twoheadrightarrow C$ with $A = \ker(p)$, and in any case is written as

$$0 \to A \to B \to C \to 0. \tag{4.2}$$

In general $\mathrm{Hom}(X, -)$ is *left exact*, i.e.,

$$0 \to \mathrm{Hom}(X, A) \to \mathrm{Hom}(X, B) \to \mathrm{Hom}(X, C).$$

A functor F which preserves surjections is called *right exact*

$$F(A) \to F(B) \to F(C) \to 0,$$

and a functor (like $\mathrm{Hom}(P, -)$ for P projective) which preserves both, is called *exact*

$$0 \to \mathrm{Hom}(P, A) \to \mathrm{Hom}(P, B) \to \mathrm{Hom}(P, C) \to 0.$$

Another name for a short exact sequence (**??**) is an *extension*. The trivial extension where $B = A \oplus C$ with the obvious inclusion i and projection p, is also known as a *split* SES. Determining when SES are split and not split is a very important topic.

For now we can make one easy observation

Lemma 4.14. *If P is projective then any SES of the form*

$$0 \to A \to B \to P \to 0$$

is split.

Proof. We apply the first definition of projective, to the situation

$$
\begin{array}{ccc}
B & \xrightarrow{\ p\ } & P \\
& \nwarrow \quad \diagdown & \uparrow {\scriptstyle 1} \\
& {\scriptstyle \widetilde{\varphi}} \diagdown & \\
& & P
\end{array}
$$

The image of $\widetilde{\varphi}$ in B must be an object complementary to the image of A in B. So B is the direct sum of these images. $\qquad\square$

Like usual in category theory, notions come in pairs. The dual notion to projective object is injective object, whose definition is obtained by reversing all arrows.

Definition 4.9. An object I is said to be *injective* if for all i and φ there exists $\widetilde{\varphi}$ as in the diagram

$$
\begin{array}{ccc}
A & \overset{i}{\hookrightarrow} & B \\
\varphi \downarrow & \swarrow\ \widetilde{\varphi} & \\
I & &
\end{array}
$$

We now also get the dual lemma

Lemma 4.15. *If I is injective then any SES of the form*

$$0 \to I \to B \to C \to 0$$

is split.

Lemma 4.16. *The irreducible objects of $\mathcal{O}$ are precisely the $\mathfrak{g}$-modules $L(\Lambda)$ for $\Lambda \in \mathfrak{h}^*$.*

Proof. Any object in $\mathcal{O}$ has at least one highest weight vector. So let v_Λ be one such for our simple object. This induces a nonzero morphism $M(\Lambda) \to M$, which image is a submodule of M and so must be all of M. So we have a surjection $M(\Lambda) \to M$. The kernel is contained in the maximal graded submodule, and if it were not equal, the image of the maximal graded submodule would be a nontrivial submodule of M. So we are done. $\qquad\square$

We noticed that the proof of projectivity of $M(\Lambda)$ works not only for $\Lambda \in P_+$, but far more generally, even for a dense set of weights Λ. I drew a picture illustrating this. TO DO: Write out this discussion, and include it and the picture here.

Definition 4.10. Let $\gamma \in \mathbb{C}$, we define the "very weak block"[1] $\mathcal{O}^{(\gamma)}$ to be the subcategory of $\mathcal{O}$ consisting of objects M which satisfy $M = M^{(\gamma)}$.

As we have seen, any object of $\mathcal{O}$ decomposes into a direct sum of objects in distinct very weak blocks. But on the other hand, each very weak block contains an uncountable number of non-isomorphic objects (in particular instance all the Verma modules $M(\lambda)$ for $|\lambda+\rho|^2 - |\rho|^2 = \gamma$).

We have seen in the discussion above that we can do better. For example let us consider $M(\lambda)$ and $M(\mu)$ in the same very weak block, i.e.,

$$|\lambda + \rho|^2 = |\mu + \rho|^2.$$

If we do not also have $\lambda - \mu \notin Q$, then $\lambda \notin \operatorname{supp} M(\mu)$ and $\mu \notin \operatorname{supp} M(\lambda)$, so there is no way for there to exist a morphism $M(\lambda) \to M(\mu)$, and vice versa. The two modules "cannot interact" in some sense.

To capture this idea in a clean way we make the following definition.

---

[1] Jethro's terminology, no one really calls it by this name.

Definition 4.11. A *composition series* of $M \in \mathcal{O}$ is a sequence of submodules

$$M = M_0 \supset M_1 \supset \ldots \supset M_n \supset M_{n+1} = 0,$$

such that each quotient $L_i = M_i/M_{i+1}$ is irreducible.

Proposition 4.17. *Each object of $\mathcal{O}$ has a composition series.*

I skipped the proof in class, but here it is in full.

Proof. Let $M = \bigoplus_{\gamma \in \mathbb{C}} M^{(\gamma)}$ be the decomposition into generalised eigenspaces of Ω. Without loss of generality we may assume $M = M^{(\gamma)}$, i.e., $\Omega - \gamma$ is locally nilpotent on M.

The idea is to start with the series

$$M = M_0 \supset M_1 = 0.$$

If M is irreducible, we are done. If not then there is a submodule S which we may insert into the series:

$$M = M_0 \supset S \supset M_2 = 0.$$

Then we continue looking for submodules of S and submodules of M_0 containing S, etc. The question is why this process should terminate in a finite number of steps.

We shall simultaneously keep track of the characters of the quotient modules. For each weight μ we have

$$\dim(M_0)_\mu \geq \dim(M_1)_\mu \geq \ldots \geq \dim(M_n)_\mu \geq 0,$$

Since $\dim(M_0)_\mu$ is finite, clearly there comes a time after which new quotients have zero μ-weight space.

We know that on a highest weight vector of weight λ the operator Ω acts as $(\lambda, \lambda + 2\rho)$. But since the Killing form is positive definite[2], the set of possible highest weights of subquotients is finite.

Note that we are not claiming that any given submodule M_i is generated by a singular vector. The following could happen

---
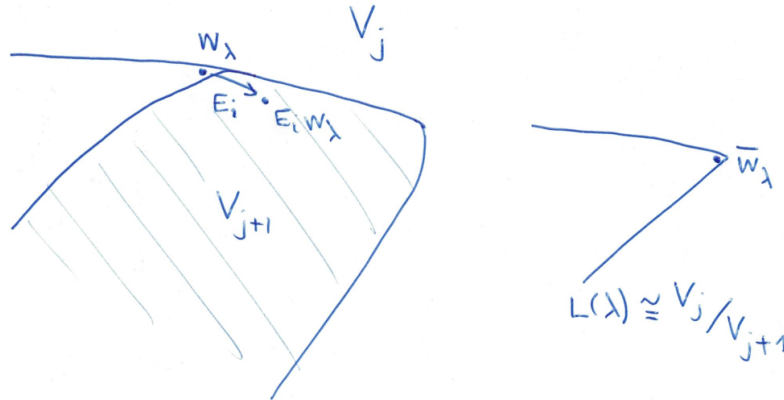
[2]Only if we work over the reals! So I am cheating a bit here, we are also using that the roots are defined over the reals.

Nevertheless we know that Ω descends to the same scalar on the quotient M_i/M_{i+1}. $\square$

Proposition 4.18. *If $L(\lambda)$ is a composition factor of $M(\Lambda)$, then $\Lambda \geq \lambda$ and*

$$|\lambda + \rho|^2 = |\Lambda + \rho|^2.$$

Proof. The first condition on λ is clear since $\operatorname{supp} M(\Lambda) = \Lambda - Q_+$, and the second follows from the fact that Ω acts in $M(\Lambda)$ by the scalar $(\Lambda, \Lambda + 2\rho)$, and thus acts by this constant in all subquotients. $\square$

Now we can define (weak) blocks.

Definition 4.12. The "weak block" $\mathcal{O}^{[\lambda]}$ consists of all objects of $\mathcal{O}$ whose composition factors are of the form $L(\lambda_i)$ where

- $\lambda_i - \lambda \in Q$,

- $|\lambda_i + \rho|^2 = |\lambda + \rho|^2$.

In other words we have specified a certain equivalence class on weights $\lambda \in \mathfrak{h}^*$, and defined weak blocks to consist of objects with composition factors $L(\lambda_i)$ where all λ_i come from a single equivalence class. By the theory we have developed so far we can prove:

Proposition 4.19. *Each object of $\mathcal{O}$ decomposes into a direct sum of objects, which all lie in distinct weak blocks.*

For the majority of weights $\lambda \in \mathfrak{h}^*$ the intersection of the sets

$$\{\mu \in \mathfrak{h}^* \mid |\mu + \rho|^2 = |\lambda + \rho|^2\}$$

and

$$\lambda + Q$$

is $\{\lambda\}$. In such a case we have $L(\lambda) = M(\lambda)$ and the only objects in $\mathcal{O}^{[\lambda]}$ are direct sums of copies of $M(\lambda)$. The other extreme, in some sense, is the case in which $\lambda \in P$. This is the case we shall generally focus on.

Remark 4.20. *Of course there is one case in which the weak block $\mathcal{O}^{[\lambda]}$ is small, even though the weight is integral, namely $\lambda = -\rho$. Then for all weights $\mu = -\rho + \alpha \in -\rho + Q$ we have*

$$|\mu + \rho|^2 = |\alpha|^2 > 0,$$

while $|-\rho + \rho|^2 = 0$. In particular $M(-\rho)$ is irreducible.

If $\lambda \in \mathfrak{h}^*$ and $w \in W$ then we can consider the following operation:

$$w \circ \lambda = w(\lambda + \rho) - \rho.$$

Since $w : \mathfrak{h}^* \to \mathfrak{h}^*$ is an orthogonal transformation (i.e., $(w(\lambda), w(\mu)) = (\lambda, \mu)$) we see that

$$|w \circ \lambda + \rho|^2 = |\lambda + \rho|^2.$$

If we further assume $\lambda \in P$ then $w \circ \lambda - \lambda \in Q$, so all the weights

$$\{L(w \circ \lambda) \mid w \in W\}$$

lie in the same weak block $\mathcal{O}^{[\lambda]}$.

In fact a stronger relation on composition factors can be proved.

Proposition 4.21. *Let $L(\lambda)$ be a composition factor of $M(\Lambda)$. Then $\Lambda \geq \lambda$ and*

$$\lambda = w(\Lambda + \rho) - \rho$$

for some $w \in W$.

The proof of this is not difficult, but it will come somewhat later in the course. It uses the *Harish-Chandra isomorphism* This result motivates the following definition.

Definition 4.13. The "block" $\mathcal{O}^\lambda$ consists of all objects of $\mathcal{O}$ whose composition factors are of the form $L(\lambda_i)$ where

- $\lambda_i - \lambda \in Q$,

- $\lambda_i = w \circ \lambda$ for some $w \in W$.

It is then possible to prove:

Proposition 4.22. *Each object of $\mathcal{O}$ decomposes into a direct sum of objects, which all lie in distinct blocks.*

5 The Weyl complete reducibility theorem

Let $\mathfrak{g}$ be a Lie algebra and A, C two $\mathfrak{g}$-modules. An extension of A by C is a short exact sequence of $\mathfrak{g}$-modules

$$0 \to A \to B \to C \to 0.$$

More precisely, two such SES are considered isomorphic if there exists an isomorphism $B \to B'$ such that

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle 1} & & \downarrow & & \downarrow{\scriptstyle 1} & & \\
0 & \longrightarrow & A & \longrightarrow & B' & \longrightarrow & C & \longrightarrow & 0
\end{array}
$$

commutes. One can actually define the category of extensions of A by C, with morphisms being any morphism $B \to B'$ making the diagram above commute. Then the 5-lemma implies that any such morphism is actually an isomorphism. In fancy language this means the category of extensions of A by C is a *groupoid*.

Any SES of vector spaces is split. Let us take an extension as above, and randomly choose a splitting at the level of vector spaces, i.e.,

$$0 \longrightarrow A \xrightarrow{\ i\ } B \underset{s}{\overset{p}{\longrightarrow}} C \longrightarrow 0$$

where $p \circ s = 1_C$. Consider the following construction: For $x \in \mathfrak{g}$ and $c \in C$ the element

$$xs(c) - s(xc)$$

lies in $\ker(p) = \mathrm{im}(i)$. Thus we have a $\mathbb{C}$-linear map

$$\mathfrak{g} \to \mathrm{Hom}(C, A)$$
$$x \mapsto (c \mapsto [x, s]c).$$

Write $M = \mathrm{Hom}(C, A)$ and ξ_B for the linear map.

You might observe that $M = \mathrm{Hom}(C, A)$ is naturally a left $\mathfrak{g}$-module, which prompts the question of whether ξ_B is a morphism of $\mathfrak{g}$-modules. In general it is not. What is the $\mathfrak{g}$-module structure of M? Well $\mathrm{Hom}(C, A) \cong C^* \otimes A$ and we know that C^* is a $\mathfrak{g}$-module via

$$[xf](c) = -f(xc),$$

and $C^* \otimes A$ is a $\mathfrak{g}$-module via

$$x(f \otimes a) = (xf) \otimes a + f \otimes (xa),$$

so that

$$[x\xi](c) = x\xi(c) - \xi(xc).$$

There is a functor from $\mathfrak{g}$-modules to vector spaces of $\mathfrak{g}$-*invariants*, denoted $M \mapsto M^{\mathfrak{g}}$, and it is clear that

$$\mathrm{Hom}_{\mathfrak{g}}(C, A) = \mathrm{Hom}(C, A)^{\mathfrak{g}}.$$

Below we shall introduce the Chevalley-Eilenberg complex of $\mathfrak{g}$, for now we just consider:

$$M \to \mathrm{Hom}(\mathfrak{g}, M) \to \mathrm{Hom}(\wedge^2 \mathfrak{g}, M) \to \cdots$$

(these are spaces of $\mathbb{C}$-linear morphisms, not $\mathfrak{g}$-linear morphisms). Let us write δ for the differentials, the first for example is (if $m \in M$)

$$\delta(m) : x \mapsto xm,$$

and the second is (if $f \in \mathrm{Hom}(\mathfrak{g}, M)$)

$$\delta(f) : x \wedge y \mapsto xf(y) - yf(x) - f([x, y]).$$

Anyway, our extension B has given us an element ξ_B of $\mathrm{Hom}(\mathfrak{g}, M)$. Actually, has it really? We had to choose a splitting s to define ξ_B. Suppose we chose a different splitting s', then we would obtain the different map

$$\xi'_B : x \mapsto (c \mapsto [x, s']c).$$

Let us consider the difference between ξ_B and ξ'_B. It will act as

$$x \mapsto (c \mapsto [x, s - s']c).$$

Observe that $p(s - s')c = c - c = 0$, so $(s - s')c \in \ker(p) = \mathrm{im}(i) = A$. Thus the difference $s - s'$ has given us a linear morphism $C \to A$, i.e., an element $\eta \in \mathrm{Hom}(C, A) = M$. Now the map

$$x \mapsto (c \mapsto [x, \eta]c)$$

is $\delta(\eta)$! So the element ξ_B is not well-defined as an element of $\mathrm{Hom}(\mathfrak{g}, M)$, but it is well-defined modulo $\mathrm{im}(\delta)$.

One also checks that $\xi_B \in \ker(\delta)$. So there is a well-defined class

$$[\xi_B] \in H^1(\mathfrak{g}, M) = \frac{\ker(\delta)}{\mathrm{im}(\delta)}.$$

It is a theorem that $B \mapsto [\xi_B]$ is an isomorphism from the isomorphism classes of extensions to $H^1(\mathfrak{g}, M)$. How is this proved? Well, $[\xi_B] = 0$ just means that there is some $\eta \in M$ such that $\xi_B = \delta(\eta)$, i.e.,

$$(x \mapsto (c \mapsto [x, s]c)) = (x \mapsto (c \mapsto [x, \eta]c))$$

or rather

$$[x, s]c = [x, \eta]c$$

for all $x \in \mathfrak{g}$ and all $c \in C$. But now we can use η to "correct" s, setting $\tilde{s} = s - \eta$. Since $\eta(c) \in A$ we have $p \circ \tilde{s} = 1$ as required, but now

$$x\tilde{s}(c) = \tilde{s}(xc),$$

so $\tilde{s}$ is a morphism of $\mathfrak{g}$-modules. Therefore vanishing of $[\xi_B]$ has been used to exhibit a splitting of the extension B, as required.

Definition 5.1. A Lie algebra $\mathfrak{g}$ is said to be semisimple if it contains no nonzero solvable ideal.

Theorem 5.1 (Cartan's criterion). *A Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is nondegenerate.*

This version (the semisimple version) of Cartan's criterion is equivalent to another version (the solvable version)

Theorem 5.2. *Let $\mathfrak{g} \subset \mathrm{End}(M)$ be a Lie subalgebra. Then $(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}])_M = 0$ if and only if $\mathfrak{g}$ is solvable.*

One direction of this follows easily from Lie's theorem. The other direction is very tricky.

Proposition 5.3 (Whitehead)**.** *Let $\mathfrak{g}$ be a semisimple Lie algebra and M a finite dimensional $\mathfrak{g}$-module. Then $H^1(\mathfrak{g}, M) = 0$.*

Proof. We will have to use Cartan's criterion (in the solvable form), and for this we need to get into a position where $\mathfrak{g} \subset \mathrm{End}(M)$. Consider the ideal $\mathfrak{g}_0 \subset \mathfrak{g}$ defined by

$$\mathfrak{g}_0 = \{x \in \mathfrak{g} \mid xM = 0\}.$$

Since $\mathfrak{g}$ is semisimple, both the ideal $\mathfrak{g}_0$ and the quotient $\overline{\mathfrak{g}} = \mathfrak{g}/\mathfrak{g}_0$ are semisimple too. Since $\mathfrak{g}_0$ acts trivially on M, we see that m acquires the structure of an $\overline{\mathfrak{g}}$-module. Furthermore the morphism $\overline{\mathfrak{g}} \to \mathrm{End}(M)$ is injective, so we have $\overline{\mathfrak{g}} \subset \mathrm{End}(M)$. So far so good. We also note that any $f \in Z^1(\mathfrak{g}, M)$ is induced by some $\overline{f} \in Z^1(\overline{\mathfrak{g}}, M)$. Why is this? Because $[\mathfrak{g}_0, \mathfrak{g}_0] = \mathfrak{g}_0$, so any element of $\mathfrak{g}_0$ can be written as a linear combination of terms $[y, z]$ for $y, z \in \mathfrak{g}_0$, and on such terms we have

$$f([y, z]) = yf(z) - zf(y) = 0 - 0 = 0.$$

So indeed we lose nothing by assuming $\mathfrak{g} \subset \mathrm{End}(M)$.

We define a Casimir operator for M. Consider the trace form $(\cdot, \cdot)_M : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ defined by

$$(x, y)_M = \mathrm{Tr}_M \, xy.$$

We claim that $()_M$ is nondegenerate. Indeed the radical of the form is an ideal in $\mathfrak{g}$, but also by Theorem **??** it is a solvable Lie algebra. Thus by semisimplicity of $\mathfrak{g}$, the radical vanishes and the form is nondegenerate. We take $\{a_i\}$ to be a basis of $\mathfrak{g}$ and $\{b_i\}$ to be the dual basis relative to $()_M$. We define the Casimir to be

$$\Omega = \Omega_M = sum_i a_i b_i.$$

Let $M = \bigoplus_{\gamma \in \mathbb{C}} M_\gamma$ be the decomposition of M into generalised eigenspaces of the Casimir operator Ω_M, and let M' be the direct sum over $\gamma \neq 0$ so that $M = M_0 \oplus M'$. If $M = M_0$ then we have

$$0 = \mathrm{Tr}_M \, \Omega = \sum_i \mathrm{Tr}_M \, a_i b_i = \sum_i (a_i, b_i)_M = \dim \mathfrak{g},$$

so in fact we must have $M' \neq 0$ and so $\dim(M_0) < \dim(M)$. This will form the basis of an inductive argument. Indeed we shall prove $H^1(\mathfrak{g}, M') = 0$ so that $H^1(\mathfrak{g}, M) = H^1(\mathfrak{g}, M_0)$.

A key observation is that Ω is invertible in M'. Let $f \in Z^1(\mathfrak{g}, M')$. By Lemma **??**

$$\Omega f(x) = xm, \quad \text{where} \quad m = \sum_i a_i f(b_i).$$

Consider the element $\Omega^{-1}m \in M'$, we shall verify that

$$f = \delta(\Omega^{-1}m),$$

so that $\in B^1(\mathfrak{g}, M')$ and $H^1(\mathfrak{g}, M') = 0$ as desired. Indeed

$$f(x) = \Omega^{-1}xm = x\Omega^{-1}m$$

and we are done (the fact that Ω^{-1} commutes with the action of x follows from the fact that Ω commutes with the action of x).

$\square$

Lemma 5.4. *Let $f : \mathfrak{g} \to M$ be a 1-cocycle, i.e., an element of $Z^1(\mathfrak{g}, M)$. Let $\{a_i\}$ and $\{b_i\}$ be dual bases of $\mathfrak{g}$ relative to an invariant bilinear form, and Ω the corresponding Casimir element. Then*

$$x\sum_i a_i f(b_i) = \Omega f(x) \qquad \text{for all } x \in \mathfrak{g}.$$

Remark 5.5. *It is easy to construct examples of non-split extensions of modules over non-semisimple Lie algebras. For example let $\mathfrak{g} = H$ the Heisenberg Lie algebra, namely $H = \langle p, q, c \rangle$ where*

$$p = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad c = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then $\mathfrak{a} = \mathbb{C}c$ is an ideal, and the inclusion $\mathfrak{a} \subset H$ gives an extension

$$0 \to \mathfrak{a} \to H \to Q \to 0,$$

which is easily checked to be nontrivial.

Remark 5.6 (Thanks Thadeu!). *Suppose $\dim(\mathfrak{g}) < \infty$, and suppose $H^1(\mathfrak{g}, M) = 0$ for all finite dimensional $\mathfrak{g}$-modules M. Is it true that $\mathfrak{g}$ must be semisimple? Well, we consider the adjoint $\mathfrak{g}$-module, it in particular must be isomorphic to a direct sum of irreducible modules. If we choose such a decomposition, it is easy to see that each factor is an ideal, each of which has no nonzero proper ideals and is therefore simple as a Lie algebra. So indeed such $\mathfrak{g}$ is semisimple.*

6   Lie algebra cohomology

This material is mostly from [**?**, Section 7.7].

Definition 6.1. Let $\mathfrak{g}$ be a Lie algebra. The Chevalley-Eilenberg complex $V_\bullet(\mathfrak{g})$ of $\mathfrak{g}$ is the following sequence of free left $U(\mathfrak{g})$-modules:

$$V_n(\mathfrak{g}) = U(\mathfrak{g}) \otimes \wedge^n \mathfrak{g},$$

with differential

$$
\begin{aligned}
d(u \otimes \xi_1 \wedge \ldots \wedge \xi_n) = & \sum_{i=1}^{n} (-1)^{i+1} (u\xi_i) \otimes \xi_1 \wedge \ldots \wedge \widehat{\xi_i} \wedge \ldots \wedge \xi_n \\
& + \sum_{i<j} (-1)^{i+j} u \otimes [\xi_i, \xi_j] \wedge \xi_1 \wedge \ldots \wedge \widehat{\xi_i} \wedge \ldots \wedge \widehat{\xi_j} \wedge \ldots \wedge \xi_n
\end{aligned}
$$

Using this complex we can give the "explict" definitions of homology and cohomology of $\mathfrak{g}$-modules (the "implicit" definitions being in terms of derived functors).

Definition 6.2. Let M be a left $U(\mathfrak{g})$-module, then the cohomology $H^n(\mathfrak{g}, M)$ of $\mathfrak{g}$ with coefficients in M is the cohomology of the complex

$$C^n = \operatorname{Hom}_{U(\mathfrak{g})}(V_n(\mathfrak{g}), M).$$

Observe that an element of $\operatorname{Hom}_{U(\mathfrak{g})}(U(\mathfrak{g}) \otimes \wedge^n \mathfrak{g}, M)$ is uniquely determined by its restriction to (a basis of) $\wedge^n \mathfrak{g}$, and so we have a natural identification

$$C^n = \operatorname{Hom}_{\mathbb{C}}(\wedge^n(\mathfrak{g}), M).$$

This is how the picture we are setting up here is related to that drawn in the previous section.

Now we can define homology. Notice that it is easier to state the definition for right $U(\mathfrak{g})$-modules.

Definition 6.3. Let N be a right $U(\mathfrak{g})$-module, then the homology $H_n(\mathfrak{g}, M)$ of $\mathfrak{g}$ with coefficients in M is the homology of the complex

$$C_n = N \otimes_{U(\mathfrak{g})} V_n(\mathfrak{g}).$$

In order to

Proposition 6.1. *The sequence $V_n(\mathfrak{g})$ with differentials defined above is indeed a complex (i.e., $d \circ d = 0$), and it is exact for all $n \neq 0$. Indeed*

$$
H_i(V_\bullet(\mathfrak{g})) = \begin{cases} \mathbb{C} & i = 0 \\ 0 & i \neq 0 \end{cases}
$$

The fact that $d \circ d = 0$ can be proved by direct calculation, and we leave it as an exercise.

To prove the statement on homology is essentially as difficult as proving the PBW theorem. In fact we will introduce a filtration, and use it to manipulate the problem to a point where we can use the result of the PBW theorem directly.

2024-01-15. Got up to here.

Definition 6.4. Let V be a vector space, an increasing filtration of V is a sequence of vector subspaces of V:

$$F_0 V \subset F_1 V \subset F_2 V \subset \ldots$$

(by convention we write $F_{-1}V = 0$). The filtration is called complete if $\bigcup_p F_p V = V$.

The associated graded is the vector space

$$\mathrm{gr}\, V = \bigoplus_{p \geq 0} F_p V / F_{p-1} V,$$

we write $\pi_p : F_p V \to \mathrm{gr}_p V$ for the quotient maps on components.

Fix an ordered basis $\{x_1, \ldots, x_N\}$ of $\mathfrak{g}$.

We introduce the PBW filtration in $U(\mathfrak{g})$ by setting

$$F_p U(\mathfrak{g}) = \langle x_{i_1} \cdots x_{i_r}) \mid r \leq p \rangle.$$

One checks that

$$\pi_p(x) \pi_q(y) = \pi_{p+q}(xy)$$

is a well-defined product in the vector space $\mathrm{gr}\, U(\mathfrak{g})$ and is commutative. The fancy statement of the PBW theorem is as follows.

Theorem 6.2 (PBW). *The canonical morphism*

$$S(\mathfrak{g}) \to \mathrm{gr}\, U(\mathfrak{g})$$

is an isomorphism.

To analyse the Chevalley-Eilenberg complex we do something similar. Define

$$F_p V_n = \langle (x_{i_1} \cdots x_{i_r}) \otimes x_{j_1} \wedge \cdots \wedge x_{j_s} \mid i_1 \leq \ldots \leq i_r,\, j_1 < \ldots < j_s \text{ and } r + s \leq p \rangle,$$

so that

$$F_0 V_n \subset F_1 V_n \subset \ldots \subset V_n$$

for each n. We easily verify that $d(F_p V_n) \subset F_p V_{n-1}$, so actually we have a filtration of complexes

$$F_0 V_\bullet \subset F_1 V_\bullet \subset \ldots \subset V_\bullet.$$

Now we set up the spectral sequence associated with the filtration. There are conventions for how to do this, but to avoid obscuring the idea I will ignore them. The idea is this: the filtered complex gives a collection of objects parametrised by two integers, namely $F_p V_n$. We arrange the associated graded objects

$$\mathrm{gr}_p(V_n) = F_p V_n / F_{p-1} V_n$$

on a $\mathbb{Z} \times \mathbb{Z}$ grid and call this "page 0" of the spectral sequence. Actually the differential $V_n \to V_{n-1}$ induces certain morphisms between the objects on page 0. At each position we have an object with a morphism going in and one coming out. We take the cohomology.

We now have a $\mathbb{Z} \times \mathbb{Z}$ grid of cohomology objects, and we call this "page 1" of the spectral sequence. The "page 2" will be obtained from page 1 again as a cohomology, page 3 from page 2, and so on to infinity. Now I will explain the morphisms in more detail.

Page 0 is not mysterious. We know that $d(F_p V_n) \subset F_p V_{n-1}$ and $d(F_{p-1} V_n) \subset F_{p-1} V_{n-1}$, so it is clear that $d : V_n \to V_{n-1}$ induces a morphism

$$d^{(0)} : \mathrm{gr}_p(V_n) \to \mathrm{gr}_p(V_{n-1}).$$

If we choose to place $\mathrm{gr}_p(V_n)$ at the position (p, n) on page 0 then $d^{(0)}$ goes vertically downward one step. We might say that $d^{(0)}$ has bidegree $(0, -1)$.

$$
\begin{array}{cccc}
\mathrm{gr}_0 V_3 & \mathrm{gr}_1 V_3 & \mathrm{gr}_2 V_3 & \mathrm{gr}_3 V_3 \\
\downarrow & \downarrow & \downarrow & \downarrow \\
\mathrm{gr}_0 V_2 & \mathrm{gr}_1 V_2 & \mathrm{gr}_2 V_2 & \mathrm{gr}_3 V_2 \\
\downarrow & \downarrow & \downarrow & \downarrow \\
\mathrm{gr}_0 V_1 & \mathrm{gr}_1 V_1 & \mathrm{gr}_2 V_1 & \mathrm{gr}_3 V_1 \\
\downarrow & \downarrow & \downarrow & \downarrow \\
\mathrm{gr}_0 V_0 & \mathrm{gr}_1 V_0 & \mathrm{gr}_2 V_0 & \mathrm{gr}_3 V_0
\end{array}
$$

The notation for page N is E^N, so in this case we would we would say that we have defined

$$E^0_{p,n} = \mathrm{gr}_p(V_n)$$

To get the differential on page 1 we consider the quotient $F_p V_n / F_{p-2} V_n$ now. It is clear that there is a SES

$$0 \to \frac{F_{p-1} V_n}{F_{p-2} V_n} \to \frac{F_p V_n}{F_{p-2} V_n} \to \frac{F_p V_n}{F_{p-1} V_n} \to 0,$$

or rather

$$0 \to \mathrm{gr}_{p-1}(V_n) \to \frac{F_p V_n}{F_{p-2} V_n} \to \mathrm{gr}_p(V_n) \to 0.$$

We have such a SES for each value of n, and it is straightforward to check that the differential $d : V_n \to V_{n-1}$ induces a map between the SES corresponding to n and $n-1$. So we actually

33

have a SES of complexes

$$0 \to \operatorname{gr}_{p-1}(V_\bullet) \to \frac{F_p V_\bullet}{F_{p-2} V_\bullet} \to \operatorname{gr}_p(V_\bullet) \to 0.$$

Passing to cohomology yields a long exact sequence, and we are particularly interested in the connecting morphisms

$$\delta : H_n(\operatorname{gr}_p(V_\bullet)) \to H_{n-1}(\operatorname{gr}_{p-1}(V_\bullet)).$$

We have $E^1_{p,n} = H_n(\operatorname{gr}_p(V_\bullet))$ by the definition of page 1. So the connecting morphism above is telling us there is a morphism

$$E^1_{p,n} \to E^1_{p-1,n-1}.$$

So page 1 looks as follows

$$
\begin{array}{cccc}
E^1_{0,3} & E^1_{1,3} & E^1_{2,3} & E^1_{3,3} \\[2em]
E^1_{0,2} & E^1_{1,2} & E^1_{2,2} & E^1_{3,2} \\[2em]
E^1_{0,1} & E^1_{1,1} & E^1_{2,1} & E^1_{3,1} \\[2em]
E^1_{0,0} & E^1_{1,0} & E^1_{2,0} & E^1_{3,0}
\end{array}
$$

The differentials now have degree $(-1,-1)$.

The general yoga of spectral sequences is that each page is cohomology of the previous page, and itself acquires a new differential. The bidegree of the differential on page N is a "linear function" of N, i.e.,

$$\deg d^{(N)} = (\alpha_0, \beta_0) + (\alpha, \beta)N.$$

In the case above

$$\deg d^{(0)} = (0,-1) \quad \text{and} \quad \deg d^{(1)} = (-1,-1),$$

so

$$(\alpha_0, \beta_0) = (0,-1) \quad \text{and} \quad (\alpha, \beta) = (-1,0).$$

The differentials on page 2 will have bidegree $(-1,-2)$, and so on.

Finally, what's the point of setting up a spectral sequence? The key theorem is that under fairly general hypotheses, "page ∞" recovers the homology of our original complex $V_\bullet$, in the following sense

$$\operatorname{gr} H_n(V_\bullet) = \bigoplus_{k \in \mathbb{Z}} E^\infty_{(0,n) + (\alpha,\beta)k}$$

But we almost never really go to page ∞. Instead, in 99% of cases we do the following: with a judicious choice of filtration, the complexes on page 1 (or occasionally page 2) become

rather easy to analyse, and in fact vanish in practically all degrees. This implies vanishing of the corresponding $H_n(V_\bullet)$.

In the present situation, the first thing to observe is that

$$E^0_{p,n} = \mathrm{gr}_p(V_n) = 0 \quad \text{whenever } n > p.$$

This is clear when you think about it: an element of V_n is a wedge product of at least n of the x_{j_k}, putting it in $F_{\geq n}V_n$. So page 0 really looks like

$$
\begin{array}{cccc}
 & & & \mathrm{gr}_3 V_3 \\
 & & & \downarrow \\
 & & \mathrm{gr}_2 V_2 & \mathrm{gr}_3 V_2 \\
 & & \downarrow & \downarrow \\
 & \mathrm{gr}_1 V_1 & \mathrm{gr}_2 V_1 & \mathrm{gr}_3 V_1 \\
 & \downarrow & \downarrow & \downarrow \\
\mathrm{gr}_0 V_0 & \mathrm{gr}_1 V_0 & \mathrm{gr}_2 V_0 & \mathrm{gr}_3 V_0
\end{array}
$$

In fact if we push this observation a bit further, we note that F_pV_p must be the span of terms of the form

$$1 \otimes x_{j_1} \wedge \cdots \wedge x_{j_p},$$

so in fact $E^0_{p,p}$ is identified with $\wedge^p\mathfrak{g}$. At the other extreme $V_0 = U(\mathfrak{g})$ and here the filtration is exactly the PBW filtration. So $E^0_{p,0} = \mathrm{gr}_p U(\mathfrak{g})$ can be identified with $S(\mathfrak{g})_p$ the vector subspace of degree p polynomials in $\mathfrak{g}$. Other terms, such as $\mathrm{gr}_2 V_1 \cong S(\mathfrak{g})_1 \otimes \wedge^1\mathfrak{g} = \mathfrak{g} \otimes \mathfrak{g}$ are tensor products of the two types of term.

$$
\begin{array}{ccc}
 & & \wedge^2\mathfrak{g} \\
 & & \downarrow \\
\mathfrak{g} & & \mathfrak{g} \otimes \mathfrak{g} \\
\downarrow & & \downarrow \\
\mathbb{C} \qquad \mathfrak{g} & & S(\mathfrak{g})_2
\end{array}
$$

What about the differentials? We see that the second sum in the definition of d vanishes in the associated graded, and the induced differential on E^0 is

$$d(u \otimes \xi_1 \wedge \ldots \wedge \xi_n) = \sum_{i=1}^{n}(-1)^{i+1}(u\xi_i) \otimes \xi_1 \wedge \ldots \wedge \widehat{\xi_i} \wedge \ldots \wedge \xi_n.$$

An observation: if we set $K_n = \bigoplus_{p\geq 0} E^0_{p,n}$ then we naturally have $K_n = S(\mathfrak{g}) \otimes \wedge^n\mathfrak{g}$ with differential defined by the formula above.

This is a definition that we can make for any commutative algebra in fact: Let R be a commutative $\mathbb{C}$-algebra, and let $\{x_1, \ldots, x_N\}$ be a set of elements in R. Write L for the linear span of the x_i, and define

$$K_n = R \otimes \wedge^n L$$

with the differential as above. This is called the Koszul complex of R relative to the set $\{x_i\}$. It is important in commutative algebra, and the general theorem about it is the following:

Definition 6.5. A sequence $x_1, \ldots, x_N$ in the commutative $\mathbb{C}$-algebra R is a "regular sequence" if each x_i is a non zero divisor in the quotient $R/(x_1, \ldots, x_{i-1})$.

Theorem 6.3. *Let $I \subset R$ denote the ideal in R generated by $\{x_1, \ldots, x_N\}$. If the set (in some order) is a regular sequence then*

$$H_i(K_\bullet) = \begin{cases} R/I & i = 0 \\ 0 & i \neq 0 \end{cases}$$

Since $S(\mathfrak{g})$ is a polynomial algebra in $x_1, \ldots, x_N$, each quotient $S(\mathfrak{g})/(x_1, \ldots, x_{i-1})$ is just a smaller polynomial algebra, none of which has any zero divisors at all. So the theorem applies, and we find

$$H_i(K_\bullet) = \begin{cases} S(\mathfrak{g})/(\mathfrak{g}) = \mathbb{C} & i = 0 \\ 0 & i \neq 0 \end{cases}$$

Thus the spectral sequence has "collapsed" at the first page with $E^s_{p,n} = E^1_{p,n}$ for all $s \geq 1$, concentrated in bidegree $(0,0)$. This proves Proposition **??**.

I claimed that the exactness of the Chevalley-Eilenberg complex was of essentially the same difficulty as the PBW theorem. We have certainly used the PBW theorem in an essential way. For honesty's sake, we should go over the proof of Theorem **??**, to show that it is not hiding some other deep and difficult result. That said, we actually only use the very special case in which R is a polynomial algebra $\mathbb{C}[x_1, \ldots, x_N]$, and our regular sequence is precisely the list of generators $x_1, \ldots, x_N$. So let's just prove this case, it's almost trivial.

First think about the case $N = 1$: the complex is

$$\mathbb{C}[x] \to \mathbb{C}[x],$$

with the differential being multiplication by x. Clearly the homology of this complex is 1 in degree 0 and 0 in all other degrees.

Now consider the case $N = 2$: the complex is

$$\mathbb{C}[x, y] \to \mathbb{C}[x, y] \oplus \mathbb{C}[x, y] \to \mathbb{C}[x, y],$$

the first map is determined by $1 \mapsto (-y, x)$, and the second is $(f, g) \mapsto xf + gy$.

In fact we deal with this by making a general definition and a standard result:

Definition 6.6. Let $C_\bullet$ and $D_\bullet$ be two homological complexes. We define the tensor product complex to be

$$(A \otimes B)_n = \bigoplus_{i+j=n} A_i \otimes B_j,$$

with differential given by

$$d(a \otimes b) = d(a) \otimes b \pm a \otimes d(b).$$

Now

Proposition 6.4. *Let A and B be complexes of vector spaces. Then*

$$H(A \otimes B) \cong H(A) \otimes H(B).$$

Now we check that

$$K(\mathbb{C}[x_1, \ldots, x_N]) = K(\mathbb{C}[x_1]) \otimes \cdots \otimes K(\mathbb{C}[x_N]),$$

and so the general statement follows from the case $N = 1$.

GOT TO HERE ON 2024-01-17

7 First interlude on sheaves: sheaves of vector spaces

As the course advances we will start using sheaves. In this section we review sheaves of vector spaces over a topological space. A particularly important example of such sheaves are the "local systems". Later we will increase the level of complexity by considering sheaves of $\mathcal{O}$-modules, and then sheaves of $\mathcal{D}$-modules. Finally we close the circle with the Riemann-Hilbert correspondence, which asserts that certain nice $\mathcal{D}$-modules (so called regular holonomic $\mathcal{D}$-modules) are essentially the same thing as local systems.

In these notes a "neighbourhood" of a point is always taken to mean an open neighbourhood of the point.

Definition 7.1. Let X be a topological space. A presheaf $\mathcal{F}$ of vector spaces on X consists of an assignment $U \mapsto \mathcal{F}(U)$ of a vector space $\mathcal{F}(U)$ to each open $U \subset X$, and a morphism of vector spaces $\mathrm{Res}_{U,V} \mathcal{F}(U) \to \mathcal{F}(V)$ to each inclusion of open sets $V \hookrightarrow U$, such that

- $\mathrm{Res}_{U,U} = \mathrm{Id}$,

- whenever $W \subset V \subset U$ we have $\mathrm{Res}_{V,W} \circ \mathrm{Res}_{U,V} = \mathrm{Res}_{U,W}$.

The presheaves on X form a category, with a morphism $\phi : \mathcal{F} \to \mathcal{G}$ consisting of a collection $\phi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ for each open U, compatible with all restrictions.

The most primitive example, and one of the most important, is the following: Let M be a fixed vector space. We assign $\mathcal{F}(U) = M$ for every open $U \subset X$, and set $\mathrm{Res}_{U,V} = \mathrm{Id}$ for all $V \subset U$. A presheaf of this form is known as a constant presheaf on X.

Another fundamental example is the presheaf C_X in which $C_X(U)$ is the vector space of continuous functions $U \to \mathbb{C}$, and the maps $\mathrm{Res}_{U,V}$ are in fact restriction of functions.

Definition 7.2 ([?, Section 1]). Let X be a topological space. A sheaf $\mathcal{F}$ of vector spaces on X consists of an assignment $x \mapsto \mathcal{F}_x$ of a vector space $\mathcal{F}_x$ to each point $x \in X$, and a topology on the disjoint union $\mathcal{F} = \bigcup_{x \in X} \mathcal{F}_x$ (we write $\pi : \mathcal{F} \to X$ for the obvious surjection) such that

- for each $x \in X$ and $f \in \mathcal{F}_x$, there exists a neighbourhood U of x in X and a neighbourhood $\widetilde{U}$ of f in $\mathcal{F}$ such that $\pi : \widetilde{U} \to U$ is a homeomorphism,

- the function $\mathcal{F} \to \mathcal{F}$ $((x,f) \mapsto (x,-f))$ is continuous,

- let $\mathcal{F} \times_X \mathcal{F} \subset \mathcal{F} \times \mathcal{F}$ stand for the subset of pairs $((x,f),(y,g))$ in which $x = y$, then the function $\mathcal{F} \times_X \mathcal{F} \to \mathcal{F}$ $((x,f,g) \mapsto (x,f+g))$ is continuous.

Let $\mathcal{F}$ be a presheaf on X. For any point $x \in X$ we define the stalk $\mathcal{F}_x$ to be the colimit (also known in this case as "direct limit")

$$\operatorname*{colim}_{x \in U} \mathcal{F}(U)$$

over all U containing x, along the morphisms Res.

direct-limit.png

{\color{red}Correct this diagram, it should read "colimit".} Elements of the stalk are called germs, and they consist of pairs (U, s) where U is an open containing x and $s \in \mathcal{F}(U)$. Two such pairs represent the same germ if there is an open subset of the intersection on which the restrictions of the sections coincide. Addition of germs (V_1, s_1) and (V_2, s_2) is carried out by restriction to $V_1 \cap V_2$ and addition of the sections s_1 and s_2 in $\mathcal{F}(V_1 \cap V_2)$.

We recall the general definition of a colimit.

Definition 7.3. Let J be a diagram in a category $\mathcal{C}$, that is, a collection of objects $A_i \in \mathcal{C}$ and a collection of morphisms $\alpha_{ij} : A_i \to A_j$. We consider the *undercategory* which consists of tuples: $X \in \mathcal{C}$ an object, and compatible $f_i : A_i \to X$ for each A_i (compatible means $f_j \circ \alpha_{ij} = f_i$ whenever $\alpha_{ij} : A_i \to A_j$ is part of J). A colimit of J is an initial object in the undercategory of J, that is, an object $(X, \{f_i\})$ such that for any other object $(X', \{f_i'\})$, there exists a unique morphism $X \to X'$ making a bunch of obvious diagrams commute.

In this level of generality the colimit of a diagram may or may not exist. In the case of the collection of vector spaces $\mathcal{F}(V)$ for $x \in V$ considered above, the colimit exists and is constructed as the vector space of germs described above.

The set of stalks of a presheaf forms a sheaf. Indeed, starting with a presheaf $\mathcal{F}$ on X we are going to consider the disjoint union $\mathcal{F}^+ = \bigcup_{x \in X} \mathcal{F}_x$, and equip it with a topology such that it becomes a sheaf. The topology is the one generated by the following collection of open sets: for $U \subset X$ open and $s \in \mathcal{F}(U)$ we define

$$[U, s] = \{(x, s(x)) \mid x \in U\} \subset \mathcal{F}^+.$$

(See [**?**, Chap. II, Exercise 1.13].) The construction $\mathcal{F} \mapsto \mathcal{F}^+$ defines a functor, denoted Ét, or simply $(-)^+$.

Given a sheaf $\mathcal{F}$ we can form a presheaf $\mathcal{F}^\circ$ (my notation) by setting

$$\mathcal{F}^\circ(U) = \{f : U \to \mathcal{F} \mid f(x) \in \mathcal{F}_x \text{ and } f \text{ is continuous}\}.$$

In fact the functors $(-)^+$ and $(-)^\circ$ are adjoint functors.

Remark 7.1. *This definition of sheaf looks quite similar to the definition of a vector bundle over X. Recall that a vector bundle of finite rank n consists of a total space E with a projection $\pi : E \to X$ and certain homeomorphisms $\phi_U : \pi^{-1}(U) \to U \times \mathbb{R}^n$ (for $\{U\}$ some sufficiently fine cover of X) such that $\phi_U \circ \phi_V^{-1}$ are appropriately continuous on restrictions. In particular for every $x \in X$ there is the* fibre *$\pi^{-1}(x) \cong \mathbb{R}^n$. On the other hand we have the set of sections*

$$\Gamma(U, E) = \{s : U \to \pi^{-1}(U) \mid s \text{ continuous and } \pi \circ s = 1_U\}.$$

These vector spaces form a presheaf (denote it $\mathcal{F}$), and we have the vector space $\mathcal{F}_x$, the stalk of $\mathcal{F}$ at x. It is good to remember that the stalk and fibre are quite different. In fact the fibre at x of our vector bundle is finite dimensional, while the stalk of the associated sheaf is the infinite dimensional vector space of "germs of sections of E" at x.

The functors $(-)^+$ and $(-)^\circ$ are adjoint. It is customary (and probably more intuitive) to work in the category of presheaves, and define a sheaf according to some conditions which are always satisfied by presheaves of the form $\mathcal{F}^\circ$. Explicitly

Definition 7.4. Let $\mathcal{F}$ be a presheaf on X. We say $\mathcal{F}$ is a sheaf if the following conditions hold:

- If $U \subset X$ is open and $\{V_i\}$ is an open cover of U, and we are given $s_i \in \mathcal{F}(V_i)$ for each i, such that for all i, j

$$\mathrm{Res}_{V_i, V_i \cap V_j}\, s_i = \mathrm{Res}_{V_j, V_i \cap V_j}\, s_j,$$

  then there exists $s \in \mathcal{F}(U)$ such that

$$s_i = \mathrm{Res}_{U, V_i}\, s$$

  for each i.

- If $s \in \mathcal{F}(U)$ and $\{V_i\}$ is a cover as above, and $\mathrm{Res}_{U, V_i}\, s = 0$ for each i, then $s = 0$. In particular the section s whose existence is assured by the last item, is unique.

The constant presheaf associated with vector space M is not a sheaf. We define the *constant sheaf* to be the sheafification of the constant presheaf, and we denote it $\underline{M}_X$. For all $x \in X$ the stalk of $\underline{M}_X$ at x is a copy of M. Again, note that this is very different than the trivial vector bundle $E = X \times M$, whose *fibre* (not stalk) at x is a copy of M.

To see why the constant sheaf $\underline{M}_X$ is different than the constant presheaf $\mathcal{F}$ defined by $\mathcal{F}(U) = M$ for all open U, we consider $U \subset X$ the union of a collection of disjoint open sets $\{U_i\}$, and a section $v_i \in \mathcal{F}(U_i)$ for each i. If the v_i are not all equal, then there is no section in $\mathcal{F}(U)$ which restricts to the various v_i. The stalk of $\mathcal{F}$ at every point is M, so $\underline{M}_X = \mathcal{F}^+$ is $X \times M$ with a certain weird topology, and the sheaf of sections $(\mathcal{F}^+)^\circ$ turns out to be the set of locally constant M-valued functions on X.

The presheaf C_X is already a sheaf, it turns out.

Another interesting example: Let M be a fixed vector space, and $x \in X$ any point. We assign $\mathcal{F}(U) = M$ for every open $U \subset X$ which contains x, and $\mathcal{F}(U) = 0$ if not. We set $\mathrm{Res}_{U,V} = \mathrm{Id}$ if V contains x and 0 otherwise. This is a sheaf, and a sheaf of this form is known as a skyscraper sheaf on X.

In [**?**] (in the exercise referenced above), the topology on $\mathcal{F}^+$ is described as being the "strongest" one for which the functions $U \to \mathcal{F}^+$ defined by $x \mapsto (x, s(x))$ are continuous, for all open $U \subset X$ and all sections $s \in \mathcal{F}(U)$. One topology is said to be "stronger" than another (or "finer" than another) if it contains all of the open sets of the latter. Thus the claim is plausible since the sets $[U, s]$ have been declared to be open, and if we were to declare any more sets to be open, it would become more difficult for a function $U \to \mathcal{F}^+$ to be continuous.

A morphism of presheaves $\varphi : \mathcal{F} \to \mathcal{G}$ consists of the data of a morphism $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ for each open U, which are compatible with the restriction maps in $\mathcal{F}$ and $\mathcal{G}$.

If we view a sheaf as a special kind of presheaf, then a morphism of sheaves is just a morphism of presheaves. If we adopt the fibre-wise definition then a morphism $\varphi : \mathcal{F} \to \mathcal{G}$ of sheaves consists of a morphism $\varphi_x : \mathcal{F}_x \to \mathcal{G}_x$ for each stalk, such that the union $\varphi : \mathcal{F} \to \mathcal{G}$ is continuous. We say that φ is injective if each stalk map is injective, and surjective if each stalk map is surjective.

Translating the notions of injectivity and surjectivity back into the language of presheaves, we eventually find that φ is injective/surjective exactly if: for every $x \in X$ there exists a neighbourhood U of x such that $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective/surjective.

Note that this definition is different than the following "naive" (and wrong) definition: a morphism of sheaves $\varphi : \mathcal{F} \to \mathcal{G}$ is injective/surjective if $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective/surjective for all open $U \subset X$. It's an exercise to show that an injective morphism of sheaves is always injective in this naive sense, but a surjective morphism is not necessarily surjective in the naive sense.

Regarding injectivity, the sheaf axiom used in an essential way. Indeed the idea is that if on some open U the map $\varphi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ were non injective, then the φ_V would have to be non injective on some strictly smaller open $V \subset U$ because of the sheaf gluing axiom.

Proposition 7.2. *Let X be a topological space. The category of all sheaves on X is an abelian category, with kernel and cokernel defined stalk-by-stalk.*

Let $\mathcal{F}$ be a sheaf on X and $U \subset X$ an open subset. We have $\mathcal{F}(U)$ which is just some vector space, and we have another entity $\mathcal{F}|_U$ a sheaf on U defined by

$$\mathcal{F}|_U(V) = \mathcal{F}(V) \quad \text{for all } V \subset U \text{ open.}$$

Definition 7.5. Let M be a finite dimensional vector space. Let $\mathcal{F}$ be a sheaf on X such that for every $x \in X$ there exists a neighbourhood U of x and an isomorphism

$$\mathcal{F}|_U \cong \underline{M}_U.$$

Then we say that $\mathcal{F}$ is a local system on X.

The constant sheaf is an obvious example of such a thing, but there are other examples too. Let $X = \mathbb{C}^\times = \mathbb{C} \backslash \{0\}$ and let
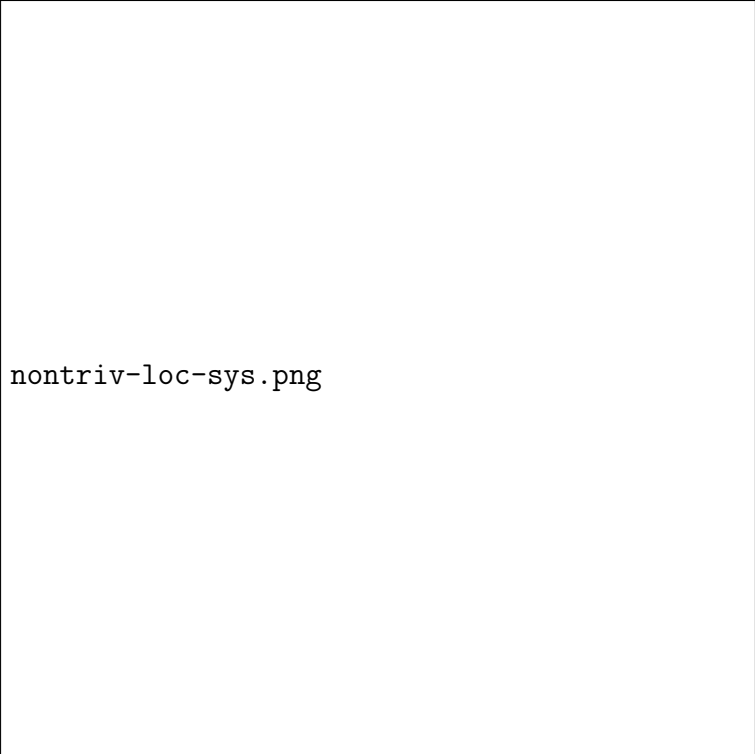
$$\mathcal{F}(U) = \{ax^{2/3} \mid a \in \mathbb{C}\}.$$

The function $f(x) = x^{2/3}$ is not globally defined without a branch cut (in which case it becomes discontinuous across the cut). But in any sufficiently small ball we can define the function unambiguously, more precisely in

$$B_\varepsilon(a) = \{x \in \mathbb{C} \mid |x - a| < \varepsilon\}$$

for $a \in \mathbb{C}^\times$ and $0 < \varepsilon < |a|$.

We obtain a sheaf which is isomorphic to the constant sheaf $\mathbb{C}_U$ when restricted to any of these balls, but which has no nonzero sections over X and so is definitely not isomorphic to $\mathbb{C}_X$.

```
nontriv-loc-sys.png
```

A comment on specifying a sheaf on a base. See [**?**, Section 6.30] for a careful discussion. We recall that a topological space consists of a set X and a set τ of subsets $U \subset X$ called the open subsets, such that τ satisfies certain axioms. By definition a *base* of τ is a subset $B \subset \tau$ such that every $U \in \tau$ is equal to the union of a (possibly infinite) collection of sets in B. The open balls form a base of the usual metric topology of $\mathbb{R}^n$ for example.

To specify a sheaf it is sufficient to specify it on a base of the topology. If one wants to be formal, it is possible to define the notion of "sheaf on B" (where B is a base of a topology), and then prove a proposition:

Proposition 7.3. *Let $\mathcal{F}$ be a sheaf on a topological space (X, τ) with base B, then the restriction of $\mathcal{F}$ to the open sets in B defines a sheaf on the base, this operation defines a functor from the category of sheaves on τ to the category of sheaves on B, and this functor is an equivalence of categories.*

8 Image functors of sheaves

Let X, Y be topological spaces, and $f : X \to Y$ a continuous function. Let $\mathcal{F}$ be a sheaf on X. It is possible to push $\mathcal{F}$ forward to Y as follows: Let $V \subset Y$ be an open set, then $f^{-1}(V) \subset X$ is open, we define a presheaf $f_*\mathcal{F}$ on Y by

$$f_*\mathcal{F}(V) := \mathcal{F}(f^{-1}(V)).$$

One verifies that this is actually a sheaf.

The pushforward is always left exact, but not in general right exact. That is, let

$$0 \to \mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to 0$$

be a short exact sequence of sheaves on X, then we claim that

$$0 \to f_*\mathcal{F}_1 \to f_*\mathcal{F}_2 \to f_*\mathcal{F}_3.$$

We have seen that injectivity/surjectivity can be tested by passing to a sufficiently small neighbourhood of each point. The thing is that no matter how small a neighbourhood V of $y \in Y$ we take, the preimage $f^{-1}(V)$ might be "big", and so surjectivity of $f_*\mathcal{F}_2(V) \to f_*\mathcal{F}_3(V)$ might fail for *all* neighbourhoods of y. The case of injectivity is different: An injection of sheaves really is injective on every open, and therefore injectivity of $\mathcal{F}_1 \to \mathcal{F}_2$ implies injectivity of $f_*\mathcal{F}_1 \to f_*\mathcal{F}_2$.

Now we describe pullback $f^{-1}\mathcal{G}$ of a sheaf $\mathcal{G}$ on Y. Since the image $f(U) \subset Y$ of open $U \subset X$ is not open in general, how to turn a presheaf on Y into a presheaf on X is not so clear. The definition, in fact, is the following: Let $\mathcal{G}$ be a presheaf on Y, and let $U \subset X$ be open. Consider the set of all open $V \subset Y$ containing $f(U)$. We form a colimit[3] over these, defining

$$f^{-1}\mathcal{G}(U) = \operatorname*{colim}_{V \supset f(U)} \mathcal{G}(V). \tag{8.1}$$

If $\mathcal{G}$ is a sheaf, then this construction is still not enough, as it happens that $f^{-1}\mathcal{G}$ might not be a sheaf. So for sheaves $\mathcal{G}$ on Y we define $f^{-1}\mathcal{G}$ to be the sheafification $(-)^+$ of the presheaf pullback described above.

The definition of pullback of a sheaf is actually easier if we use the stalk-wise definition. Then $\mathcal{G}$ has a stalk $\mathcal{G}_y$ at each $y \in Y$ and we define the sheaf $f^{-1}\mathcal{G}$ to have stalk at x given by $(f^{-1}\mathcal{G})_x = \mathcal{G}_{f(x)}$.

It's an exercise to verify that the two definitions of the pullback of a sheaf coincide. We note that the pullback functor f^{-1} is exact. This is transparent from the stalk-wise definition. See [**?**, Lemma 17.3.3(c)].

It turns out (f^{-1}, f_*) forms a left-right adjoint pair between categories of sheaves. Recall this means there are natural isomorphisms

$$\operatorname{Hom}(f^{-1}\mathcal{G}, \mathcal{F}) \cong \operatorname{Hom}(\mathcal{G}, f_*\mathcal{F}).$$

If we unpack what the two sides mean, the isomorphism becomes clear. This is a good exercise.

Here is a little more explanation. A morphism $\mathcal{G} \to f_*\mathcal{F}$ consists of a compatible system of morphisms of vector spaces (indexed by open $V \subset Y$):

$$\mathcal{G}(V) \to f_*\mathcal{F}(V) = \mathcal{F}(f^{-1}(V)).$$

---

[3]Remember: Initial object which receives maps from all the guys

Now we wish to describe morphisms $f^{-1}\mathcal{G} \to \mathcal{F}$. Since we have defined $f^{-1}(\mathcal{G})$ as a sheafification, we should comment on this before starting.

Remark 8.1 (Thanks Diego!). *The sheafification functor $(-)^+$, and the functor $(-)^\circ$ sending a sheaf to its presheaf of sections (or rather, the functor of forgetting that the presheaf is a sheaf) is itself an example of an adjoint pair. Namely for any presheaf $\mathcal{F}$ and sheaf $\mathcal{G}$ on X, we have*

$$\mathrm{Hom}(\mathcal{F}^+, \mathcal{G}) \cong \mathrm{Hom}(\mathcal{F}, \mathcal{G}^\circ).$$

This means that to specify a morphism $f^{-1}\mathcal{G} \to \mathcal{F}$ of sheaves is really the same thing as to specify a morphism $f^{-1}\mathcal{G} \to \mathcal{F}$ where this time $f^{-1}\mathcal{G}$ stands for the presheaf (**??**). Such a morphism consists of a compatible system of morphisms of vector spaces (indexed by open $U \subset X$):

$$\operatorname*{colim}_{V \supset f(U)} \mathcal{G}(V) \to \mathcal{F}(U).$$

Let us restrict to a fixed U for a moment. By definition of colimit, a morphism as above is the same thing as a collection of compatible morphisms

$$\mathcal{G}(V) \to \mathcal{F}(U)$$

running over all open $V \supset f(U)$. Switching perspective and viewing V as fixed, this is the data of a compatible system of morphisms running over all U such that $f(U)$ is contained in V. But $f(U) \subset V$ is the same thing as saying $U \subset f^{-1}(V)$. Such a system is completely and uniquely determined by the morphism $\mathcal{G}(V) \to \mathcal{F}(f^{-1}(V))$. So we have proved the adjointness property. Moreover we have seen that at its core the adjointness of (f^{-1}, f_*) hinges on the change of perspective between all open V containing $f(U)$, and all open U contained in $f^{-1}(V)$.

We reserve the symbol $\Gamma(X, -)$ for $\pi_*(-)$ the pushforward along the surjection $\pi : X \to *$ to a point. In other words $\Gamma(X, \mathcal{F})$ is $\mathcal{F}(X)$. As for pushforwards in general, the functor

$$\Gamma(X, -) : \mathrm{Sh}(X) \to \mathbf{Vec}$$

is left-exact. If we could show that $\mathrm{Sh}(X)$ has enough injectives (which it does), then we could compute derived functors $R^n\Gamma(X, -)$.

Derived functors of pushforward to a point have a very natural interpretation. If X is a sufficiently nice topological space they recover singular cohomology of X.

Theorem 8.2. *Let X be a topological manifold, then there are isomorphisms*

$$R^n\Gamma(X, \mathbb{R}_X) \cong H^n_{sing}(X).$$

The theorem is true for topological spaces much more general than topological manifolds. It is true (for example) if X is *paracompact* and *locally contractible*.

It's actually not hard to show that $\mathrm{Sh}(X)$ has enough injectives, using a construction due to Godement.

Let X be a topological space and let $\{M_x\}$ be a completely arbitrary collection of vector spaces indexed by $x \in X$. We define a presheaf $\mathcal{M}$ by

$$\mathcal{M}(U) = \prod_{x \in U} M_x$$

(so an element of this vector space is an arbitrary choice of $v_x \in M_x$ for each $x \in U$). This is easily seen to be a sheaf. One notes that the stalk $\mathcal{M}_x$ of this sheaf at x is very different from M_x.

Indeed consider the example in which $X = \mathbb{R}$ and $M_x = 0$ unless x is of the form $1/n$ for $n \in \mathbb{Z}_{\geq 1}$, in which case $M_x = \mathbb{C}$. Then the stalk at 0 is the vector space of "tails", namely the vector space of sequences

$$(a_1, a_2, a_3, a_4, \ldots)$$

modulo the equivalence relation $\sim$ that says

$$(a_1, a_2, a_3, a_4, \ldots) \sim (b_1, b_2, b_3, b_4, \ldots)$$

if there exists $N \geq 1$ such that $a_n = b_n$ for all $n \geq N$.

Given a sheaf $\mathcal{F}$, one defines the Godement "envelope" of $\mathcal{F}$ as the above construction applied to the collection of vector spaces $\mathcal{F}_x$. There is a canonical morphism of sheaves

$$\mathcal{F} \to \mathrm{Gode}(\mathcal{F}),$$

indeed a section $s \in \mathcal{F}(U)$ induces $s_x \in \mathcal{F}_x$ at each $x \in U$, and we combine these to define a section $\widehat{s} \in \mathrm{Gode}(\mathcal{F})(U)$.

The construction described above (and so in particular the Godement envelope) is an example of an injective object in the category $\mathrm{Sh}(X)$. Indeed let $i : \mathcal{F} \to \mathcal{G}$ be an injective morphism of sheaves, and let $\mathcal{M}$ be a sheaf of the form described above: $\mathcal{M}(U) = \prod_{x \in U} M_x$. Let a morphism $f : \mathcal{F} \to \mathcal{M}$ be given. By definition this consists of a collection of compatible morphisms

$$\mathcal{F}(U) \to \mathcal{M}(U) = \prod_{x \in U} M_x,$$

i.e., a collection of compatible morphisms

$$\mathcal{F}(U) \to M_x$$

where U runs over all open sets containing x. By definition of stalk this is the same thing as a morphism of vector spaces $\mathcal{F}_x \to M_x$ (note: M_x here, not $\mathcal{M}_x$). Since vector spaces are always injective objects, we can extend this morphism along the injection $\mathcal{F}_x \to \mathcal{G}_x$ to obtain $\mathcal{G}_x \to M_x$ for each x. It remains to confirm that these morphisms define a morphism of sheaves $\mathcal{G} \to \mathcal{M}$ making

$$\begin{array}{ccc} \mathcal{F} & \overset{i}{\hookrightarrow} & \mathcal{G} \\ {\scriptstyle f}\downarrow & \swarrow{\scriptstyle \widetilde{f}} & \\ \mathcal{M} & & \end{array}$$

Remark 8.3. *The world's shortest introduction to derived functors. The general narrative of derived functors starts from the notions of left and right exact functors. If the abelian category $\mathcal{C}$ has enough projectives then any object $A \in \mathcal{C}$ can be given a projective resolution $P_\bullet \to A$, and any such resolution can be used to compute the values on A of the left derived functors $L_i F$ of a right exact functor F (example: tensor product). Similarly if F is left exact (example: $F(-) = \Gamma(X, -)$ from sheaves on X to* **Vec***), then an injective resolution $A \to I^\bullet$ can be used to compute the values on A of the right derived functors $R^i F$.*

GOT TO HERE ON 2024-01-22

Although it is important to know that enough projectives or enough injectives exist, in the day-to-day business of computing derived functors it is often possible to get away with using less exotic objects. Given a left exact functor F, there is a notion of F-injective object, and it is a theorem that a resolution by F-injective objects suffices to compute derived functors $R^i F$.

In the case of sheaves, one has the following notion:

Definition 8.1. A sheaf $\mathcal{F}$ on X is said to be flabby if the restriction morphism $\mathcal{F}(X) \to \mathcal{F}(U)$ is surjective for every open set $U \subset X$.

It is then a standard theorem (I found it in these notes [**?**, Proposition 6.9], look up a textbook reference)

Theorem 8.4. *A flabby sheaf is $\Gamma(X, -)$-injective, indeed $\Gamma(X, -)$-acyclic (which is stronger).*

The sheaves of singular cochains are not necessarily injective, but if X is a topological manifold then they are flabby. So by the homological algebra general nonsense above, these sheaves can be used to resolve if one is computing global sections.

The proof of Theorem **??** then proceeds by introducing sheaves of singular cochains. Roughly speaking: for $n \in \mathbb{Z}_{\geq 0}$ a basic singular chain in $U \subset X$ is a continuous function $\sigma : \Delta^n \to U$, where Δ^n is the standard n-simplex, and a singular chain is a formal linear combination of finitely many basic ones. The vector space of singular chains is denoted $C_n(U)$ and singular cochains are elements of $C^n(U) = C_n(U)^*$ the dual vector space. These form presheaves, which we then sheafify to obtain sheaves $\mathcal{C}^n_{\text{sing}}$. The usual coboundary produces a complex

$$0 \to \mathcal{C}^0_{\text{sing}} \to \mathcal{C}^1_{\text{sing}} \to \mathcal{C}^2_{\text{sing}} \to \cdots$$

of sheaves which resolves $\underline{\mathbb{R}}_X$ (this is proved by passing to small open sets in X which are sufficiently nice to check explicitly). A key point is that the sheaves $\mathcal{C}^n_{\text{sing}}$ are flabby, so the derived global sections of $\underline{\mathbb{R}}_X$ are computed as the cohomology of the complex of vector spaces

$$0 \to \Gamma(X, \mathcal{C}^0_{\text{sing}}) \to \Gamma(X, \mathcal{C}^1_{\text{sing}}) \to \Gamma(X, \mathcal{C}^2_{\text{sing}}) \to \cdots$$

Now these vector spaces are exactly the vector spaces of singular cochains which are used to define singular cohomology in basic topology.

9 Weyl character formula

Where we left off our study of category $\mathcal{O}$, we had seen that every $M \in \mathcal{O}$ possesses a composition series with factors a finite set of irreducible modules $L(\lambda_i)$, each with a finite multiplicity $b_i \in \mathbb{Z}_{\geq 0}$. These factors and multiplicities are independent of the specific composition series, so we write $[M : L(\lambda)]$ for b_i. This induces an equality of characters

$$\mathrm{Ch}_M = \sum_{\lambda_i} m_i \, \mathrm{Ch}_{L(\lambda_i)} \, .$$

We are particularly interested in the multiplicities

$$b_{\Lambda,\lambda} = [M(\Lambda) : L(\lambda)]$$

(where $\Lambda \in P$ for simplicity). We know that $L(\Lambda)$ appears as the first quotient $M(\Lambda)/J$, so $b_{\Lambda,\Lambda} = 1$. Then any other nonzero $b_{\Lambda,\lambda}$ occur for $\lambda < \Lambda$. Let us choose a linear order on P consistent with the partial order $\leq$. The set of coefficients

$$b_{\Lambda,\lambda} = [M(\Lambda) : L(\lambda)]$$

can be viewed as a $\mathbb{Z} \times \mathbb{Z}$ matrix, which is lower triangular with ones on the diagonal. Hence the matrix can be inverted and we may write

$$\mathrm{Ch}_{L(\Lambda)} = \sum_{\lambda \leq \Lambda} m_{\Lambda,\lambda} \, \mathrm{Ch}_{M(\lambda)} \, . \tag{9.1}$$

We have $m_{\Lambda,\Lambda} = 1$ and $m_{\Lambda,\lambda} \in \mathbb{Z}$.

We will prove a formula for the character of $L(\Lambda)$ for $\Lambda \in P_+$.

Combining the formula for the character of the Verma with equation (**??**) and multiplying through by e^ρ yields

$$R \, \mathrm{Ch}_{L(\Lambda)} = \sum_{\lambda \leq \Lambda} m_{\Lambda,\lambda} e^\lambda, \qquad R = \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha}). \tag{9.2}$$

For reasons that will become clear soon, we will multiply through to obtain:

$$e^\rho R \, \mathrm{Ch}_{L(\Lambda)} = \sum_{\lambda \leq \Lambda} m_{\Lambda,\lambda} e^{\lambda + \rho}. \tag{9.3}$$

First we will show that $e^\rho R$ is W-skew invariant by checking it on the generating reflections r_i. We need the following two facts:

- Δ is W-invariant.

- The set $\Delta_+ \backslash \{\alpha_i\}$ is invariant under r_i.

The first fact is part of the classical theory we reviewed in the first lessons.

As for the second fact: write $\beta \in \Delta_+ \backslash \{\alpha_i\}$ as a $\mathbb{Z}_+$-linear combination of the α_k with a positive coefficient for some simple root $\alpha_j \neq \alpha_i$. Then $r_i(\beta)$ also has a positive coefficient of α_j and hence lies in $\Delta_+ \backslash \{\alpha_i\}$, because every root is either positive or negative.

Now we calculate

$$r_i(e^\rho R) = e^{\rho - \alpha_i}(1 - e^{\alpha_i}) \prod_{\alpha \in \Delta_+ \backslash \{\alpha_i\}} (1 - e^{-\alpha}) = -e^\rho R,$$

which is the W-skew invariance asserted above.

We have assumed that $\Lambda \in P_+$. We now show that this implies $\mathrm{Ch}_{L(\Lambda)}$ is W-invariant.

Definition 9.1. Let $\mathfrak{g}$ be semisimple. A weight module for $\mathfrak{g}$ is said to be *integrable* if E_α acts locally nilpotently in M for all $\alpha \in \Delta$.

Let M be integrable and let $\mu \in \mathrm{supp}(M)$. For some simple root α, consider the sl_2-subalgebra $\mathfrak{a} = < E_\alpha, \alpha^\vee, E_{-\alpha} >$ and its representation in $L = \bigoplus_{n \in \mathbb{Z}} M_{\lambda + n\alpha}$. Because of integrability, each vector in L lies in a finite dimensional $\mathfrak{a}$-submodule. By an argument similar to the one we did for $\mathfrak{g}$ itself, we may deduce that each weight space of L is finite dimensional and

$$\dim(M_\lambda) = \dim(M_{r_i(\lambda)}).$$

The W-invariance of Ch_M follows.

Lemma 9.1. *If $\Lambda \in P_+$ then $L(\Lambda)$ is integrable.*

Proof. We know that $\mathfrak{n}_+$ acts locally nilpotently, since we checked $L(\Lambda)$ is in $\mathcal{O}$. Now check that local nilpotence of $\mathfrak{n}_-$ holds on the highest weight vector v_Λ. The idea now is to prove local nilpotence on other vectors inductively. $\square$

An immediate corollary of W-invariance of the character is finite dimensionality of $L(\Lambda)$, which we asserted without proof in Theorem **??**.

We now see that

$$\sum_{\lambda \leq \Lambda} m_{\Lambda, \lambda} e^{\lambda + \rho}$$

must be W-skew invariant.

Consider a weight λ such that $m_{\Lambda, \lambda} \neq 0$. If $w(\lambda + \rho) = \mu + \rho$ then $m_{\Lambda, \lambda} = \epsilon(w) m_{\Lambda, \mu}$ because of skew invariance. Because $m_{\Lambda, \lambda} \neq 0$ only for $\lambda \leq \Lambda$ we deduce that $w(\lambda + \rho) \leq \Lambda + \rho$ for all $w \in W$.

Next we will show that the range of summation in (**??**) is $\lambda \in W \circ \Lambda$, where $\circ$ is the *shifted action* of the Weyl group $w \circ \lambda = w(\lambda + \rho) - \rho$. In fact it follows from Proposition **??**. We now give a proof for $\Lambda \in P_+$ (the only case we need here).

Proposition 9.2. *Let $L(\lambda)$ be a composition factor of $M(\Lambda)$, where $\Lambda \in P_+$. Then $\lambda \geq \Lambda$ and*

$$\lambda = w(\Lambda + \rho) - \rho$$

for some $w \in W$.

Proof. Let $\mu = w \circ \lambda$ be such that $\mathrm{ht}(\Lambda - \mu)$ takes its minimal value as w ranges over W. We will prove that $\mu = \Lambda$. First we will prove that $\mu \in P_+$. Indeed if $\langle \mu + \rho, \alpha_i^\vee \rangle < 0$ for some i then

$$\Lambda - r_i \circ \mu = \Lambda + \rho - r_i(\mu + \rho) = \Lambda - \mu + \langle \mu + \rho, \alpha_i^\vee \rangle \alpha_i$$

has strictly lower height than $\Lambda - \mu$. Hence $\langle \mu + \rho, \alpha_i^\vee \rangle \geq 0$ for each i, which is to say

$$\mu + \rho \in P_+.$$

Now, $\Lambda + \rho$ and $\mu + \rho$ are in P_+ and $(\Lambda + \rho) - (\mu + \rho) = \sum_i k_i \alpha_i$ where all $k_i \geq 0$. Also $|\Lambda + \rho|^2 = |\mu + \rho|^2$, which implies

$$0 = |\Lambda + \rho|^2 - |\mu + \rho|^2 = (\Lambda + \mu + 2\rho, \Lambda - \mu) = \sum_i k_i \langle \Lambda + \mu + 2\rho, \alpha_i^\vee \rangle (\alpha_i, \alpha_i).$$

The terms in the summation on the right are all positive, so each $k_i = 0$ in fact, hence $\mu = \Lambda$.

It follows that $\lambda \in W \circ \Lambda$, as claimed. $\qquad\square$

We have shown that

$$R \operatorname{Ch} L(\Lambda) = \sum_{\lambda \in W \circ \Lambda} m_{\Lambda, \lambda} e^\lambda.$$

There is one more step to the final version:

Theorem 9.3 (Weyl-Kac). *Let $\Lambda \in P_+$, then*

$$R \operatorname{Ch} L(\Lambda) = \sum_{w \in W} \epsilon(w) e^{w \circ \Lambda}. \tag{9.4}$$

To complete the proof we just need to show that the stabiliser of Λ in $(W, \circ)$ is $\{1\}$.

Proposition 9.4. *Let $\Lambda \in P_+$, then the stabiliser of Λ under the dot action of W is trivial.*

The proof relies on an annoyingly complicated lemma about the geometry of the W-action on $\mathfrak{h}^*$. But the lemma will be needed later on too, so we should take it seriously.

Lemma 9.5. *If $r_{i_1} \cdots r_{i_t}(\alpha_i) < 0$ then there exists s $(1 \leq s \leq t)$ such that*

$$r_{i_1} \cdots r_{i_s} \cdots r_{i_t} r_i = r_{i_1} \cdots r_{i_{s-1}} r_{i_{s+1}} \cdots r_{i_t}.$$

Proof. For each $k < t$ let $\beta_k = r_{i_{k+1}} \cdots r_{i_t}(\alpha_i)$, and let $\beta_t = \alpha_i$. By hypothesis $\beta_0 < 0$, and $\beta_t > 0$.

Let $\beta_{s-1} < 0$, $\beta_s > 0$, we have $\beta_{s-1} = r_{i_s}\beta_s$. Now, if $\beta \in \Delta_+$, then $(\beta + \mathbb{Z}\alpha_k) \cap \Delta \subseteq \Delta_+$ unless $\beta = \alpha_i$ simply because every root is either positive or negative. Therefore if $\beta \in \Delta_+$ and $r_i(\beta) < 0$ then $\beta = \alpha_i$. Applying this to the situation at hand shows that $\beta_s = \alpha_{i_s}$. Also $\alpha_{i_s} = w(\alpha_i)$ where $w = r_{i_{s+1}} \cdots r_{i_t}$.

The subspaces $\mathbb{R}\alpha_i$ and $\alpha_i^\perp$ are complementary, and r_i is the unique linear transformation that sends α_i to $-\alpha_i$ and restricts to the identity on $\alpha_i^\perp$. The same goes for α_{i_s}. Because w is an orthogonal transformation sending α_i to α_{i_s}, it sends $\alpha_i^\perp$ to $\alpha_{i_s}^\perp$. Therefore $r_{i_s} = w r_i w^{-1}$.

Substituting this for r_{i_s} into $r_{i_1} \cdots r_{i_s} \cdots r_{i_t} r_i$ yields $r_{i_1} \cdots r_{i_{s-1}} r_{i_{s+1}} \cdots r_{i_t}$ as desired. $\qquad\square$

Proof of Proposition **??**. Now, A is an integer matrix, so we could have defined $\mathfrak{h}_\mathbb{R}$ and $\mathfrak{h}_\mathbb{R}^*$, using $\mathbb{R}$ in place of $\mathbb{C}$. We can embed $\mathfrak{h}_\mathbb{R} \hookrightarrow \mathfrak{h}$. The action of W may be defined on $\mathfrak{h}_\mathbb{R}^*$ since the definition makes no reference to the base field.

Let
$$\mathfrak{h}_\mathbb{R}^\circ = \{h \in \mathfrak{h}_\mathbb{R} | \langle \alpha, h \rangle \neq 0 \text{ for each } \alpha \in \Delta\}.$$

The *positive (open) Weyl chamber* is the set
$$C = \{h \in \mathfrak{h}_\mathbb{R} | \langle \alpha_i, h \rangle > 0 \text{ for each } i\}.$$

We claim that C is a fundamental domain for the action of W on $\mathfrak{h}_\mathbb{R}^\circ$, i.e., for all $h \in C$, $w(h) = h$ if an only if $w = 1$. The claim above about P_+ will follow.

Suppose now that $h \in C$ and $h' = w(h) \in C$ also. Let $w = r_{i_1} \cdots r_{i_t}$ be a *reduced expression* for $w \neq e$, meaning that $t \in \mathbb{Z}_{>0}$ is minimal among all possible such representations of w. We have $\langle \alpha_{i_t}, h \rangle > 0$ by hypothesis, therefore $\langle w(\alpha_{i_t}), w(h) \rangle = \langle w(\alpha_{i_t}), h' \rangle > 0$ too.

On the other hand we claim that $w(\alpha_{i_t}) \in \Delta_-$. For assume it were the case that $w(\alpha_{i_t}) \in \Delta_+$, then we would have
$$r_{i_1} \cdots r_{i_{t-1}}(\alpha_{i_t}) = w r_{i_t}(\alpha_{i_t}) = w(-\alpha_{i_t}) \in \Delta_-,$$

and it would follow from Lemma **??** that our expression for w was not reduced, a contradiction.

Therefore $\langle w(\alpha_{i_t}), h' \rangle < 0$, which contradicts $h' = h$. So $w = e$. $\qquad\square$

The proof contains a result which it will be useful to state on its own.

Lemma 9.6. *Let $w \in W$ of length $\ell(w) = k$. Then there exists i such that wr_i has length $k - 1$. Furthermore*
$$w(\alpha_i) \in \Delta_- \qquad \text{and} \qquad (wr_i)(\alpha_i) \in \Delta_+.$$

Exercise: Write out a proof.

GOT TO HERE ON 2024-01-24

10 Category $\mathcal{O}$ has enough projective

Let us finally prove the following classification of indecomposable objects in a block of category $\mathcal{O}$ for sl_2.

Proposition 10.1. *Let* $\mathfrak{g} = sl_2$. *There are exactly* 5 *indecomposable objects in the block* $\mathcal{O}_0$, *and they are the following:*

- *The Verma* $M = M(0)$ *and its irreducible quotient* $L = L(0)$,

- *The dual Verma* $M^\vee$,

- *The irreducible Verma* $M^- = M(-2\varpi)$ *(which is self-dual),*

- *The module* V *constructed in the first exercise list as an extension*

$$0 \to M \to V \to M^- \to 0.$$

Proof. Let $A \in \mathcal{O}_0$ and consider the submodule $B = U(\mathfrak{g})A_0 \subset A$. We claim that B is isomorphic to a direct sum of copies of M and L. Indeed let $v_1 \in A_0$ be nonzero and $B^1 = U(\mathfrak{g})v_1 \subset B$.

Since v_1 is a singular vector, there is a surjection $M \to B^1$ and so B^1 is either isomorphic to M or else to its unique nontrivial quotient L.

Now grab another vector $v_2 \in A_0$ and consider the submodule B^2 generated by v_1 and v_2. (We could consider the submodule generated by v_2 alone, but that might have nontrivial intersection with B^1, and things would quickly get complicated). We can say that $B^2 \supset B^1$ with both B^1 and B^2/B^1 isomorphic to L or to M. In other words we have an extension

$$0 \to B^1 \to B^2 \to B^2/B^1 \to 0$$

which is of the following four:

$$0 \to L \to E \to L \to 0$$
$$0 \to M \to E \to M \to 0$$
$$0 \to L \to E \to M \to 0$$
$$0 \to M \to E \to L \to 0.$$

Since L is finite dimensional, we know that there are no nontrivial extensions of L by L, so in that case B^2 would be a direct sum of two copies of L.

The middle two cases are also easy to treat. In fact we have a general lemma:

Lemma 10.2. *Let*

$$0 \to A \to E \to M(\lambda) \to 0$$

be an extension, where

$$\mathrm{supp}(A) \subset \mathrm{supp}(M(\lambda)).$$

Then the extension is split.

Proof. Let $e_\lambda \in E_\lambda$ be a preimage in E of the generating vector $v_\lambda \in M(\lambda)$. Because of the condition on supports, e_λ is a singular vector in E and so by the universal property, there exists a morphism $\alpha : M(\lambda) \to E$ sending $\alpha(v_\lambda) = e_\lambda$. The image of α is some quotient of $M(\lambda)$. But since $e_\lambda \mapsto v_\lambda$ in the exact sequence, the image of α must surject to $M(\lambda)$ and so $Im(\alpha) \subset E$ is a copy of $M(\lambda)$. We have just described a splitting of the exact sequence. $\square$

What about the final case? Now there is no general result to appeal to, but we can analyse this specific case as follows. We know that

$$0 \to N \to M \to L \to 0,$$

where N in this case is the Verma module M^-. Observe that $\operatorname{Hom}(N, M) \cong \mathbb{C}$ since the weight spaces are all 1-dimensional.

Now we will use this SES to get a LES for Ext (Derived functors of Homming to M), namely

$$
\begin{array}{ccc}
\operatorname{Ext}(N,M) \xleftarrow{\quad} \operatorname{Ext}(M,M) \xleftarrow{\quad} \operatorname{Ext}(L,M) \\[2em]
\operatorname{Hom}(N,M) \xleftarrow{\quad} \operatorname{Hom}(M,M) \xleftarrow{\quad} \operatorname{Hom}(L,M) \xleftarrow{\quad} 0
\end{array}
$$

What we know so far is that

$$
\begin{array}{ccc}
\xleftarrow{\quad} 0 \xleftarrow{\quad} (?) \\[2em]
\mathbb{C} \xleftarrow{\cong} \mathbb{C} \xleftarrow{\quad} 0 \xleftarrow{\quad} 0
\end{array}
$$

and exactness of the LES gives us $\operatorname{Ext}(L, M) = 0$.

So now we know in all cases that B^2 is a direct sum of copies of L and/or M. We continue inductively (using that Ext spaces are additivie in their arguments). So $B \subset A$ is a direct sum of copies of L and M.

Consider now $Q = A/B$. The submodule C of Q generated by Q_{-2} must be isomorphic to a direct sum of copies of M^- (same reasoning as above, but simpler this time). In fact $Q = C$ as there are no further irreducible modules in $\mathcal{O}_0$ beyond L and M^- to form its composition series.

Now we either have A isomorphic to a direct sum of copies of L, M and M^-, or else A contains a nontrivial extension

$$0 \to L \to U \to M^- \to 0.$$

or

$$0 \to M \to V \to M^- \to 0.$$

There is a unique up to isomorphism extension U, and it is actually $M^\vee$. This can be proved by hand, similarly to the exercise proving uniqueness of the extension V.

Now we make a key observation: As we have proved in a previous lecture, M is a projective module, and so $M^\vee$ is an injective module. This means that if $M^\vee$ occurs as a submodule of A then it occurs as a summand. So we can consider the other summand and proceed inductively.

In fact if we could show that V is injective then we would be finished. In fact it is, as we show in the next proposition. $\qquad\square$

Proposition 10.3. *The extension V is injective.*

Proof. First we note that $V^\vee \cong V$, so it suffices to show that V is projective. Indeed consider these pictures

INSERT

To prove projectivity, we will reduce it to a general result using a clever realisation of V. Consider the Verma module $M(-\rho) = M(-\varpi)$. This module has the special property of being the unique Verma module in its block $\mathcal{O}_{-\rho}$. Since $L(-\rho) = M(-\rho)/J$ where the composition factors of J must have strictly lower weight, in fact $J = 0$ and $M(-\rho)$ is irreducible.

Because there are no Verma modules of weight higher than $-\rho$ in the same block either, we conclude that $M(-\rho)$ is projective.

Now we consider $M(-\varpi) \otimes L(\varpi)$, it is isomorphic to V! Since $L(\varpi)$ is finite dimensional, it follows that V is projective by Proposition **??** below. $\qquad\square$

Proposition 10.4. *If $P \in \mathcal{O}$ is projective, and L is finite dimensional, then $P \otimes L$ is projective.*

Proof. Recall that projectivity of P means (by definition, essentially) that the functor

$$X \mapsto \mathrm{Hom}(P, X)$$

is exact. But now

$$\begin{aligned} X \mapsto {} & \mathrm{Hom}(P \otimes L, X) \\ & \cong \mathrm{Hom}(P, \mathrm{Hom}(L, X)) \\ & \cong \mathrm{Hom}(P, X \otimes L^*) \end{aligned}$$

is exact too. $\qquad\square$

Now let us discuss embeddings of Verma modules, tying up with the notion of reduced expressions in the Weyl group.

Proposition 10.5. *Let $w = r_{i_t} \cdots r_{i_1}$ be a reduced expression, and write $w^{(k)} = r_{i_k} \cdots r_{i_1}$ (so $w^{(0)} = e$, $w^{(1)} = r_{i_1}$ and $w^{(t)} = w$). Suppose that $\Lambda \in P_+$. Then there exists an embedding*

$$M(w^{(k)} \circ \Lambda) \to M(w^{(k-1)} \circ \Lambda)$$

for each $k \geq 1$.

Proof. We wish to show that

$$m = \left\langle w^{(k-1)} \circ \Lambda, \alpha_{i_k}^{\vee} \right\rangle \in \mathbb{Z}_{\geq 0},$$

for it then follows that

$$F_{i_k}^{m+1} v_{w^{(k-1)} \circ \Lambda}$$

is a singular vector.

But we compute

$$
\begin{aligned}
\left\langle w^{(k-1)} \circ \Lambda + \rho, \alpha_{i_k}^{\vee} \right\rangle &= \left\langle w^{(k-1)}(\Lambda + \rho), \alpha_{i_k}^{\vee} \right\rangle \\
&= \left\langle \Lambda + \rho, (w^{(k-1)})^{-1}(\alpha_{i_k}^{\vee}) \right\rangle \\
&= \left\langle \Lambda + \rho, r_{i_1} \cdots r_{i_{k-1}}(\alpha_{i_k}^{\vee}) \right\rangle \\
&= \left\langle \Lambda + \rho, \beta \right\rangle
\end{aligned}
$$

where $\beta \in \Delta_+$, by Lemma **??**. The expression is therefore positive, and we are done.

The fact that $M(w^{(k)} \circ \Lambda) \to M(w^{(k-1)} \circ \Lambda)$ is injective is the following lemma. $\square$

I DREW A BUNCH OF PICTURES OF INCLUSIONS OF VERMAS FOR SL3. DRAW THEM AND INCLUDE THEM HERE.

From Lemma **??** and Proposition **??** we start to perceive the surprising importance of reduced expressions in the Weyl group W of a simple Lie algebra $\mathfrak{g}$. In particular we give the symbol ℓ to the function

$$\ell : W \to \mathbb{Z}_{\geq 0}$$

sending $w \in W$ to the length of any reduced expression of w. The length function has other combinatorial interpretations. Later we shall see:

Proposition 10.6. *For all $w \in W$ one has*

$$\ell(w) = |w(\Delta_+) \cap \Delta_-|.$$

From this it follows immediately that $0 \leq \ell(w) \leq |\Delta_+|$ for all $w \in W$. In fact the maximal value of $|\Delta_+|$ is realised by, it turns out, a unique element in the Weyl group. This element is denoted $w_\circ$ and is simply called "the longest element" of the Weyl group.

Another curious element of W is the so-called *Coxeter element*. This is the product of all the simple reflections, taken in any random order, each exactly once. It turns out this element is well-defined up to conjugacy. Its order is a number

$$h$$

called the *Coxeter number* of $\mathfrak{g}$. For example $h(sl_n) = n$ and $h(E_8) = 30$.

Another important number associated with $\mathfrak{g}$ is the *dual Coxeter number* $h^\vee$, which is defined as follows: let $\theta \in \Delta_+$ be the highest root (the unique root such that no root is higher than it) and let κ denote the Killing form of $\mathfrak{g}$ (which we restrict to $\mathfrak{h}$, use to identify $\mathfrak{h} \to \mathfrak{h}^*$, and finally transfer to a bilinear form κ on $\mathfrak{h}^*$), then

$$\kappa(\theta, \theta) = \frac{1}{h^\vee}.$$

It is very common to define

$$(\lambda, \mu) = 2h^\vee \kappa(\lambda, \mu),$$

so that $(\theta, \theta) = 2$.

Some weird relations satisfied by h and $h^\vee$:

$$h = \frac{\dim(\mathfrak{g})}{rank(\mathfrak{g})}$$

and

$$|\rho|^2 = \frac{h^\vee \dim(\mathfrak{g})}{24}.$$

GOT TO HERE ON 2024-01-25

Finally we will use Proposition **??** to show that $\mathcal{O}$ has enough projectives. For this we will need a lemma which will be useful on its own too.

Lemma 10.7 ([**?**, Theorem 3.6]). *Let $\lambda \in \mathfrak{h}^*$ and let L be a finite dimensional $\mathfrak{g}$-module. Then the tensor product*

$$M(\lambda) \otimes L$$

carries a filtration, whose associated graded is a direct sum

$$\bigoplus_\nu M(\lambda + \nu)$$

of Verma modules, where ν runs over the weights of L counted with multiplicity.

Proof. Consider the isomorphism

$$(U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} W) \otimes L \cong U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} (W \otimes L),$$

here on the left we have a tensor product of $U(\mathfrak{g})$-modules, and on the right we are considering L as a $U(\mathfrak{b})$-module by restriction. The case of interest is $W = \mathbb{C}v_\lambda$, for which the left hand side is $M(\lambda) \otimes L$.

Since W and L are weight modules, so is $W \otimes L$. We choose a basis $\{v_1, \ldots, v_n\}$ ordered so that $i \leq j$ whenever $\mathrm{wt}(v_i) \leq \mathrm{wt}(v_j)$ (so v_n is the highest of the weights). Since the action of $\mathfrak{b}$ only raises weights, the span of the subset of vectors $\{v_k, \ldots, v_n\}$ is a $\mathfrak{b}$-submodule of $W \otimes L$ for each k. In other words we have a filtration

$$W \otimes L = N_1 \supset N_2 \supset \ldots \supset N_n \supset N_{n+1} = 0,$$

with N_k being the span of $\{v_k, \ldots, v_n\}$. The associated graded is a direct sum of one-dimensional weight modules $\mathbb{C}v_{\lambda+\nu}$ where ν runs over the weights of L.

Since the functor

$$U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} (-)$$

is exact, the $\mathfrak{g}$-module

$$U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} (W \otimes L)$$

carries a filtration whose associated graded is a direct sum of modules $M(\lambda + \nu)$. This completes the proof. $\square$

The following proposition is Theorem 3.8 of the book [?] of Humphreys.

Theorem 10.8. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then the category $\mathcal{O}$ has enough projectives and enough injectives.*

Proof. We shall prove that $\mathcal{O}$ has enough projectives, that is, that for every $M \in \mathcal{O}$ there exists a projective object P and a surjection $P \to M$. By considering contragredient objects we will conclude that $\mathcal{O}$ has enough injectives too.

Let $M(\lambda)$ be a Verma module. There exists some $\Lambda = \lambda + N\rho$ for which $M(\Lambda)$ is projective, so consider $M(\Lambda) \otimes L(N\rho)$ which is also projective since $L(N\rho)$ is finite dimensional. We take an ordered basis of $L(N\rho)$ with weights increasing, as recommended in the lemma above. The first vector lies in weight $-N\rho = w_\circ(N\rho)$. The lemma then tells us that the first quotient in the filtration of $M(\Lambda) \otimes L(N\rho)$ is a copy of $M(\Lambda - N\rho) = M(\lambda)$. Thus we have proved that $M(\lambda)$ receives a surjection from the projective object $M(\Lambda) \otimes L(N\rho)$.

Since every irreducible in $\mathcal{O}$ is $L(\lambda)$ for some λ, it follws that all irreducibles receive a surjection from a projective. By induction on the length of a composition series it is possible to extend this result to all objects of $\mathcal{O}$. See the proof in [?] for details, or better yet, reconstruct them for yourself. $\square$

Let us finish up this section with a few other useful facts about Verma modules which do not need any special theoretical machinery to prove.

Lemma 10.9. *Any nonzero morphism $M(\mu) \to M(\lambda)$ is injective.*

Proof. As $U(\mathfrak{n}_-)$-modules we have $M(\mu) \cong M(\lambda)$, and $M(\mu) \to M(\lambda)$ is of the form $u \mapsto uu_0$ for some (nonzero) $u_0 \in U(\mathfrak{n}_-)$. But the ring $U(\mathfrak{n}_-)$ has no zero divisors, so $uu_0 = 0$ only if $u = 0$ and we are done. $\qquad \square$

Lemma 10.10. *For any weights λ, μ we have*

$$\dim \operatorname{Hom}(M(\mu), M(\lambda)) \leq 1.$$

Proof. First we claim that any Verma module possesses a unique irreducible submodule. In fact the ring $U(\mathfrak{n}_-)$ is left Noetherian, and the statement follows from a general argument about Noetherian rings. Namely: if $x \neq 0$ is a non-right-zero-divisor and I a left ideal in the left Noetherian ring R, then Rx meets I nontrivially. Indeed the chain of ideals

$$I \subset I + Ix \subset I + Ix + Ix^2 \subset \dots$$

must stabilise, so forsome N we have

$$x^N = a_0 + a_1 x + \dots + a_{N-1} x^{N-1}, \qquad a_i \in I.$$

Assume $I \cap Rx = 0$. Then from the equation we have $a_0 = 0$. But then we can cancel x from the equation and also deduce $a_1 = 0$, and continue in this way until we find $x = 0$, which is a contradiction. So in fact Rx has nonzero intersection with every left ideal in R. It follows, in fact, that every pair of nonzero left ideals has nonzero intersection.

Applied to $R = U(\mathfrak{n}_-)$: Suppose there were two distinct irreducible submodules of $M(\lambda) \cong U(\mathfrak{n}_-)$, corresponding to distinct left ideals I_1 and I_2. They have nonzero intersection, but the intersection is a $\mathfrak{g}$-submodule (Think about why!) and so both submodules must be equal to it, and to each other.

Now we turn to the proof of the lemma. Let f_1, f_2 be two such morphisms. Each of them sends the unique irreducible submodule $S \subset M(\mu)$ to some (irreducible) submodules S_1, S_2 of $M(\lambda)$. Again by uniqueness, these images coincide, and so there exists some constant c for which $f_1 = cf_2$. But then $g = f_1 - cf_2$ by construction satisfies $g(S) = 0$, and so by the previous lemma, $g = 0$. $\qquad \square$

11 Lie groups and algebraic groups

A Lie group is a smooth manifold G which is simultaneously a group in the sense that multiplication $G \times G \to G$ and inversion $G \to G$ are both smooth maps. An algebraic group is a smooth algebraic variety which is a group in a similar sense.

Some basic examples of Lie groups are

$$GL_n(\mathbb{R}) = \{A \in \operatorname{Mat}_{n \times n}(\mathbb{R}) \mid \det(A) \neq 0\}$$

and

$$SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \det(A) = 1\},$$

and
$$SO_n(\mathbb{R}) = \{A \in SL_n(\mathbb{R}) \mid A^T A = I_n\}.$$
We can actually take the entries of the matrices in any field, and we get a group. But having taken $\mathbb{R}$ as the field, these groups are also examples of smooth manifolds. Indeed $GL_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$ the complement of a set defined by the polynomial equation $\det(A) = 0$. For each point of $GL_n(\mathbb{R})$, there is a sufficiently small open ball centred at the point, disjoint from the $\det = 0$ set. These balls form an atlas.

The group multiplication $G \times G \to G$ is smooth since it is polynomial. The inversion map $G \to G$ is also smooth, because it is rational.

Now $SL_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$ too, but this time it is the locus cut out by the polynomial equation $\det = 1$. Why is this a manifold? We need to show every point of $SL_n(\mathbb{R})$ has a neighbourhood homeomorphic to a ball in $\mathbb{R}^{n^2-1}$ and the transition maps are smooth.

IN CLASS I SKIPPED THIS BLUE TEXT.

This is a standard sort of thing to check in differential geometry, it's an application of the inverse function theorem. Let's recall how it goes. Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a smooth function. Let $a \in U$ and write $b = f(a)$. Recall the differential
$$df_a = [\frac{\partial f}{\partial x_1}(a), \cdots , \frac{\partial f}{\partial x_n}(a)].$$
In general the differential of a smooth map $\mathbb{R}^n \to \mathbb{R}^m$ is the $m \times n$ matrix
$$df_a = [\frac{\partial f_i}{\partial x_j}(a)].$$
Suppose $df_a \neq 0$. Without loss of generality the first component of df_A is nonzero. Define
$$F : U \to \mathbb{R}^n$$
$$F(x_1, \ldots, x_n) = (f(x_1, \ldots, x_n), x_2, \ldots, x_n).$$
The differential of F at a is
$$dF_a = \begin{bmatrix} \partial_1 f(a) & \partial_2 f(a) & \partial_3 f(a) & \cdots & \partial_n f(a) \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$
Since this matrix is invertible at a, there exist neighbourhoods of $a \in U \subset \mathbb{R}^n$ and $F(a) = (b, a_2, \ldots, a_n) \in \mathbb{R}^n$ between which F is a bijection with a smooth inverse, which we might call G. The intersection of $f^{-1}(b)$ with this neighbourhood is then the graph of the function
$$g(x_2, \ldots, x_n) = G(b, x_2, \ldots, x_n).$$
The coordinates $(x_2, \ldots, x_n)$ are then a chart on this graph.

In fact this is the main idea in the proof of the following fundamental theorem in differential geometry:

Theorem 11.1 (The submersion theorem). *Let $f : M \to N$ be a submersive map of smooth manifolds, and $q \in N$. Then $f^{-1}(q) \subset M$ is an embedded submanifold.*

How do we check that $\det : \mathbb{R}^{n^2} \to \mathbb{R}$ is submersive? Well, it isn't really, at the matrix of all zeros the differential of det is also 0, so not submersive. This is not a problem, as we are working "away" from 0, but it makes it more awkward to apply the theorem.

At $a = I_n$ it is obvious that $df \neq 0$: we can just think about the family

$$A(t) = \begin{bmatrix} t & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

which satisfies $\det A(t) = t$. By smoothness there exists a ball around I_n in which $\det \neq 0$ and $d(\det) \neq 0$.

Two notions of vector field. Let X be an n-dimensional smooth manifold and let $C = C_X^\infty$ denote the sheaf of smooth functions on X, and $R = \Gamma(X, C)$ the ring of globally defined functions. Then we can consider a Lie algebra

$$\mathrm{Der}(R) = \{D : R \to R \mid D(fg) = fD(g) + gD(f)\}.$$

We can also choose a point $x \in X$, write C_x for the germ of C, and consider the ideal

$$C_x \ni \mathfrak{m}_x = \{s_x \mid s_x(x) = 0\}.$$

Consider the vector space $\mathfrak{m}_x/\mathfrak{m}_x^2$. One can check that it is n-dimensional. Then the dual space

$$T_x X = (\mathfrak{m}_x/\mathfrak{m}_x^2)^*$$

is to be thought of as the tangent space of X at x, and $\mathrm{Der}(R)$ is to be thought of as the vector space of global vector fields on X. Since we mentally picture a vector field as consisting of a vector sitting at each point of X, we ought to have, for every $x \in X$, a map

$$\mathrm{Der}(R) \to (\mathfrak{m}_x/\mathfrak{m}_x^2)^*,$$

or rather, each $D \in \mathrm{Der}(R)$ ought to give as a linear map

$$\mathfrak{m}_x/\mathfrak{m}_x^2 \to \mathbb{R}.$$

What is this map? It is the following: For $\alpha \in \mathfrak{m}_x/\mathfrak{m}_x^2$, choose a representative $(U, f_0) \in C_x$ of α, extend f_0 to a global function $f \in R$ that coincides with f_0 in U (to be able to do this we use the idea of partition of unity and bump functions), now consider $D(f)$ and in particular its image in C_x. It might be that the germ of $D(f)$ is no longer in $\mathfrak{m}_x$, but it turns out that the germ of $D(f)$ is well-defined modulo $\mathfrak{m}_x$. But $C_x/\mathfrak{m}_x \cong \mathbb{R}$, so we have just used D to produce a well-defined (and linear) map $\mathfrak{m}_x/\mathfrak{m}_x^2 \to \mathbb{R}$ as required.

Remark 11.2. *According to a mathoverflow post [?], if you try doing the $\mathfrak{m}_x/\mathfrak{m}_x^2$ construction with continuous functions you will get 0. Indeed define $f_+ = \min\{f, 0\}$ and $f_- = -\max\{f, 0\}$ so that $f = f_+ - f_-$ with $f_\pm$ non-negative at all points. We have continuous functions $g_\pm = \sqrt{f_\pm}$, and $f = g_+^2 - g_-^2 \in \mathfrak{m}_x^2$.*

The need for bump functions here might seem a symptom of an insufficiently subtle definition of vector field. Clearly, it would seem, we should define a *sheaf T_X* of vector fields on X, so that the statement above becomes the following: for each $U \subset X$ open and $x \in U$, there is a linear map

$$T_X(U) \to (\mathfrak{m}_x/\mathfrak{m}_x^2)^*.$$

We can try defining the sheaf T_X of vector fields on X by

$$T_X(U) = \mathrm{Der}(C_X^\infty(U))$$

for each open set $U \subset X$. Now what are the restriction maps $T_X(U) \to T_X(V)$ for open subsets $V \subset U$? Given a function $f \in C_X^\infty(V)$ and $D \in \mathrm{Der}(C_X^\infty(U))$ we would like to say what $D(f)$ should be. We have the same problem as above; unless we know how to extend a function from V to U then we do not know how to restrict a derivation from U to V!

If you look at the definition of T_X in an algebraic geometry text, [?, p. 180] for example, you will see it defined as

$$T_X = \underline{\mathrm{Hom}}_{\mathcal{O}_X}(\Omega_X, \mathcal{O}_X).$$

Here $\mathcal{O}_X$ is the sheaf of functions (think of it as the algebraic geometry analogue of C_X^∞), and Ω_X is another sheaf called the sheaf of 1-forms.

I will discuss Ω_X below, for now I want to focus on the notion of

$$\underline{\mathrm{Hom}}(-, -) : \mathrm{Sh}(X) \times \mathrm{Sh}(X) \to \mathrm{Sh}(X).$$

What is the construction? We could imagine assigning

$$U \mapsto \mathrm{Hom}(\mathcal{F}(U), \mathcal{G}(U)),$$

but we have no reason to believe these vector spaces possess well-defined restriction maps, any more than derivations did.

The actual definition of $\underline{\mathrm{Hom}}$ is as the sheaf defined by:

$$U \mapsto \mathrm{Hom}(\mathcal{F}|_U, \mathcal{G}|_U).$$

We must remember the crucial difference between the vector space $\mathcal{F}(U)$ and the sheaf $\mathcal{F}|_U$. The definition feels a little circular, but it is not. Essentially, a section of the sheaf $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})$ over U is a specification of how to transform from $\mathcal{F}$ to $\mathcal{G}$, not only sections over U, but also all sections over all open subsets of U. Similarly, a section of T_X over U is a specification, not only of a derivation of $C_X^\infty(U)$, but also of a derivation of the space of functions on V for each open subset V of U, and all of these derivations compatible with each other under restriction of fuctions. See also the notes [?].

Since I brought up Ω_X, I might as well define it, although we will not use it until later. Let R a commutative $\mathbb{R}$-algebra. Then

$$\Omega_R = R \otimes_{\mathbb{R}} R/I$$

where I is the vector subspace spanned by all elements of $R \otimes R$ of the form

$$a \otimes fg - af \otimes g - ag \otimes f.$$

If we start denoting $a \otimes f$ as $a\,df$, the idea becomes clear. So elements of Ω_X are differential 1-forms. The good thing is that these gadgets restrict as we expect them to, i.e., if $V \subset U$ then we have the restriction map $\mathcal{O}_X(U) \to \mathcal{O}_X(V)$. Since clearly in general a homomorphism $R \to S$ of $\mathbb{R}$-algebras induces a morphism $\Omega_R \to \Omega_S$, so we obtain natural restriction maps which aloow us to define a sheaf (or at least a presheaf, which we can then sheafify). I leave it as an exercise to confirm the universal property that Ω_R satisfies, namely that an R-linear morphism $\Omega_R \to R$ is the same thing as an element of $\mathrm{Der}(R)$.

Anyway, back to Lie groups. Let G be a Lie group and write $\mathfrak{g} = T_e G$. We are going to show that $\mathfrak{g}$ has the natural structure of a Lie algebra. For each $g \in G$ there is a smooth function

$$L_g : G \to G$$

acting as $x \mapsto gx$. Since L_g is invertible, the pullback L_g^* and pushforward L_{g*} yield inverse isomorphisms between spaces of germs of functions at x and at gx, and thence isomorphisms between tangent spaces $T_x G$ and $T_{gx} G$. For any two points $x, y \in G$ there is a unique g such that $y = gx$, and we can ask of a section ξ of TG that its values at x and y correspond under the isomorphism of tangent spaces. Such a vector field will be called *left invariant*.

Let us describe left invariance at the level of global vector fields. We have our ring R of smooth functions on G and an automorphism α_g sending f to f_1 defined by $f_1(x) = f(g^{-1}x)$ (i.e., α_g sends a bump function concentrated on x to a bump function concentrated on gx). The action of L_{g*} on vector fields corresponds to the conjugation

$$\mathrm{Der}(R) \to \mathrm{Der}(R), \qquad D \mapsto \alpha_g \circ D \circ \alpha_g^{-1}.$$

We have $\xi \in \mathrm{Der}(R)$ left invariant now if

$$\alpha_g \circ D \circ \alpha_g^{-1} = \xi.$$

An important basic fact is that the commutator of two left invariant vector fields is again left invariant. In fact it is clear that if ξ_1 and ξ_2 satisfy this condition, then

$$\xi_1 \circ \xi_2 - \xi_2 \circ \xi_1$$

does too.

Each left invariant vector field is determined by its value at e, so we think of the commutator bracket on these vector fields as an operation defined on $\mathfrak{g}$: writing ξ_a for the left invariant vector field associated with $a \in \mathfrak{g}$, we define $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$\xi_{[a,b]} = \xi_a \circ \xi_b - \xi_b \circ \xi_a.$$

61

One can check that this bracket makes $\mathfrak{g}$ a Lie algebra.

————

Construction: Let $\mathfrak{g}$ be a semisimple Lie algebra and V a faithful finite dimensional $\mathfrak{g}$-module. Consider automorphisms

$$\exp(aE_\alpha)$$

for $a \in \mathbb{C}$ and $\alpha \in \Delta$ (these expressions are polynomial in a). The subgroup $G \subset GL(V)$ is an algebraic group whose Lie algebra is isomorphic to $\mathfrak{g}$.

Exercise: Consider this construction for SL_2. Find the diagonal matrices as products of upper and lower triangular matrices.

12 The flag variety of $GL_n(\mathbb{C})$

In this section we do a few calculations with the common flag variety. Later we study flag varieties from a more abstract and general perspective.

Let V be a vector space of dimension n over a field, which can be general, but for us will always be $\mathbb{C}$.

Definition 12.1. A complete flag in V is a sequence of vector subspaces

$$V_1 \subset V_2 \subset \ldots \subset V_{n-1} \subset V,$$

where $\dim V_j = j$. The flag variety X is the set of all such flags.

As suggested by its name, the set of all flags in V can be given the structure of an algebraic variety.

First we observe that $GL(V)$ acts on V, moving subspaces around, and sending flags to flags. So X is a set with an action of the group $G = GL(V)$. The action is transitive. Indeed, given two flags F_1 and F_2, choose bases adapted to each, i.e., if F_1 is the flag

$$V_1 \subset V_2 \subset \ldots \subset V_{n-1} \subset V_n = V$$

then take $\{e_1\}$ a basis of V_1, then take e_2 such that $\{e_1, e_2\}$ is a basis of V_2, and so on. Do the same for the flag F_2, obtaining basis $\{e_1', e_2', \ldots, e_n'\}$. Finally take any element $g \in GL(V)$ which sends e_i to e_i'. By elementary group theory, then, there is a set bijection

$$X = G/B$$

where B is the stabiliser subgroup of any particular fixed flag. Let's choose a basis $\{e_1, \ldots, e_n\}$ of V and consider the "standard flag" F_0 given by $V_j = \oplus_{i \leq j} \mathbb{C}e_i$. So now $V = \mathbb{C}^n$ and $G = GL(V) = GL_n(\mathbb{C})$. A matrix fixing the standard flag must: (1) send e_1 to a nonzero multiple of e_1, (2) send e_2 to a nonzero multiple of e_2 plus any multiple of e_1, etc. Clearly $B = Stab(F_0)$ consists of the upper triangular matrices.

Remark: the restricted action of $SL(V)$ is already transitive on X, so we could have done all of the above with $G = SL_n(\mathbb{C})$ and $B \subset G$ the set of upper triangular matrices with determinant 1.

In fact this picture extends to a class of groups called *semisimple linear algebraic groups* in general. There is a notion of Borel subgroup of such a group, which we shall define below, and then it is possible to define the *flag variety* $X = G/B$ where $B \subset G$ is any Borel subgroup. An important fact is the following (see the book of Humphreys [**?**, Theorem 21.3]).

Theorem 12.1. *Let G be a simple algebraic group and $B \subset G$ a Borel subgroup. Then G/B is projective.*

One often sees the terms parabolic subgroup and parabolic subalgebra appear in this theory. Here is the definition.

Definition 12.2. A subgroup $P \subset G$ is said to be parabolic if G/P is a projective variety.

So we have the purely linguistic corollary

Corollary 12.2. *Every Borel subgroup is parabolic.*

Before we get into the nuts and bolts of linear algebraic groups, it is worthwhile to sketch some other important ingredients of the theory in the case of $G = GL_n(\mathbb{C})$.

Let $T \subset G$ be the subset of all diagonal matrices. This is isomorphic to

$$\mathbb{C}^\times \times \mathbb{C}^\times \times \cdots \times \mathbb{C}^\times$$

(with n factors), and in particular is abelian. It is called a maximal torus of G and in the general picture, the maximal tori in G correspond to the Cartan subalgebras in the corresponding Lie algebra $\mathfrak{g}$ of G.

We have introduced

$$B = \{ \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix} \}$$

already. Now we introduce also the subgroups of G

$$U = \{ \begin{pmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{pmatrix} \} \quad \text{and} \quad U^- = \{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{pmatrix} \}.$$

Consider the normaliser
$$N_G(T) = \{g \in G \mid gTg^{-1} = T\}.$$
We have $T \subset N_G(T)$ as a normal subgroup. The quotient, rather beautifully, is a finite group isomorphic to the Weyl group of $(\mathfrak{g}, \Delta)$. Since we are treating only $GL_n(\mathbb{C})$ for now, we want to check that
$$W = N_G(T)/T \cong S_n.$$
Let us say that a monomial matrix is a matrix (a_{ij}) such that for each i there is exactly one j for which $a_{ij} \neq 0$ holds. Clearly such a matrix is expressible as a permutation matrix times an invertible diagonal matrix. So we need only show the following

Proposition 12.3. *If $g \in G$ conjugates T to T then g is a monomial matrix.*

Proof. Since $X \mapsto gXg^{-1}$ is a linear operation on X, it suffices to check on a linear basis of diagonal matrices (even a basis consisting of not necessarily diagonal matrices).

So we calculate the two sides in
$$g \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{pmatrix} g,$$
say, and equate coefficients to discover that either $g_{11} = 0$, or if not then $g_{1j} = 0$ for $j = 2, \dots, n$.

We then proceed in a similar manner inductively, showing that in every row of g there is at most one (and hence exactly one) nonzero entry. $\square$

For each $w \in W$ we choose a representative $n_w \in N_G(T)$ and consider $Bn_wB \subset G$. Clearly the set does not depend on the choise of representative n_w of w, so we denote the set BwB. These are called *double cosets*. The nice thing about double cosets is that they give a partition of G, i.e.,

Theorem 12.4. *There is a decomposition as a disjoint union*
$$G = \bigcup_{w \in W} BwB.$$

Proof. First we shall show that every element $g \in G$ lies in one of the double cosets. Then we shall show their intersections are empty.

For the first part, we must show that for every $g \in G$ there exist matrices $b, b' \in B$ such that bgb' is a monomial matrix. But let us consider what left and right multiplication by elements of B mean in terms of row/column operations. The operation
$$g \mapsto g \begin{pmatrix} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

corresponds to the operation that takes α times the first column and adds it to the third column. In general, by similar such operations, we can add any multiple of column i to column j for $i < j$. Meanwhile the operation

$$g \mapsto \begin{pmatrix} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} g$$

corresponds to adding α times the third row to the first row. In general, by similar such operations, we can add any multiple of row i to row j for $i > j$. It is not hard to see that by the Gaussian process with the restricted class of row and column operations permitted to us, we are able to reduce g to a monomial matrix.

To prove the second part, we suppose that

$$Bn_w B \ni g \in Bn_{w'} B$$

so that in particular there exist $b, b' \in B$ such that

$$b' = n_w^{-1} b n_{w'}.$$

The key observation is that for each nonzero entry of $n_w^{-1} n_{w'}$ the entry of $n_w^{-1} b n_{w'}$ in the same position is also nonzero. It is not difficult to see why, and part of the justification is that all the diagonal entries of b are nonzero. It follows from the observation that $n_w^{-1} n_{w'}$ is upper triangular and hence diagonal in fact. So $w' = w$. $\qquad\square$

For each element $w \in W$ we define

$$U(w) = U \cap (n_w U^- n_w^{-1}).$$

Proposition 12.5 ([?, Corollary 23.60]). *Every element of BwB can be written in a unique way as $un_w b$ where $b \in B$ and $u \in U(w)$.*

Let's just comment on the non-uniqueness of representation in BwB. Indeed in the map

$$B \times B \to BwB, \qquad (b, b') \mapsto bn_w b'$$

the preimage of a point in general is quite large. For example for $w = e$, we have $BwB = B$ and there are many distinct ways to write a given $b_0 \in B$ as $b_0 = bb'$.

Proof. Let's also introduce $U'(w) = U \cap (n_w U n_w^{-1})$. The Lie algebra of $U(w)$ is spanned by all the $\mathfrak{g}_\alpha$ for which $\alpha \in \Delta_+$ and $w^{-1}(\alpha) \in -\Delta_+$. Similarly the Lie algebra of $U'(w)$ is spanned by all the $\mathfrak{g}_\alpha$ for which $\alpha \in \Delta_+$ and $w^{-1}(\alpha) \in \Delta_+$. It follows that $U(w)U'(w) = U$ and so $U(w)U'(w)T = B$.

The key observation is that $U'(w)n_w \subset n_w U$ IN FULTON-HARRIS THIS IS STATED AS EQUALITY, WHICH CANNOT POSSIBLY BE TRUE. We also use $T n_w = n_w T$. Now we compute

$$
\begin{aligned}
B n_w B &= U(w) U'(w) T n_w B \\
&= U(w) U'(w) n_w T B \\
&\subset U(w) n_w U T B \\
&= U(w) n_w B.
\end{aligned}
$$

That shows every element of BwB has a decomposition of the desired form. What about uniqueness? Similarly to the key observation: $n_w^{-1} U(w) n_w = U^-$. So if $n_w = u n_w b$ then $n_w^{-1} u n_w \in B$. Since $U^- \cap B = \{e\}$ we are done. $\qquad\square$

Clearly the dimension of $U(w)$ is the cardinality of $\Delta_+ \cap w^{-1}(-\Delta_+)$. According to Proposition **??** (which we have not yet proved ACTUALLY I THINK I FORGOT TO PROVE THIS) this coincides with $\ell(w)$ the length of a reduced expression for w. Furthermore, as algebraic varieties

$$
U(w) \cong \mathbb{C}^{\ell(w)}.
$$

Since we have $G = \bigcup_{w \in W} U(w) w B$, we have the corresponding partition of $X = G/B$ as

$$
X = \bigcup_{w \in W} X_w \qquad X_w = U(w) w B / B \cong U(w).
$$

The varieties X_w are called *Bruhat cells*, their closures $\overline{X}_w$ are called *Schubert varieties*. While the $X_w \cong U(w)$ are smooth (indeed they are the simplest possible kind of variety, isomorphic to affine space), the Schubert varieties are sometimes singular.

Let's do the example of $G = SL_2(\mathbb{C})$ in detail. we have

$$
B = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}, \qquad U^- = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}.
$$

We have $W = \{e, \sigma\}$, and we may take

$$
n_e = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad n_\sigma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
$$

(note that $n_\sigma^{-1} = -n_\sigma$).

Firstly $U(e) = \{I_2\}$ and so X_e consists of the single point $I_2 B / B$.

Next we examine $U(\sigma)$, we find

$$
U(\sigma) = \{ \begin{pmatrix} 1 & u \\ 0 & 0 \end{pmatrix} \} \quad \text{and} \quad X_\sigma = \{ \begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix} \}.
$$

It turns out that $G/B \cong \mathbb{P}^1$. We clearly have an identification of X_σ with $\mathbb{C}$, now let's understand how X_e is identified with the point $\{\infty\}$.

We multiply the element of X_σ parametrised by u (and we shall soon send $u \to \infty$) by a cleverly chosen element of B, which will vary with u in a way such as to compensate and make the product tend to I_2. Observe:

$$\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u^{-1} & 1 \\ 0 & u \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ u^{-1} & 1 \end{pmatrix}.$$

Clearly then in the limit $u \to \infty$ this tends to I_2.

13   Linear algebraic groups

Now we will develop a bit of the theory of linear algebraic groups over $\mathbb{C}$. The goal will be to explain an algebraic proof that G/B is a projective variety.

Let R be a finitely generated $\mathbb{C}$-algebra. Essentially by definition this means there exists a surjective homomorphism $\mathbb{C}[x_1, \ldots, x_n] \to R$, the kernel of which is an ideal I. The spectrum $X = \mathrm{Spec}(R)$ is a geometric object, we think of it either as the set of prime ideals $p \subset R$ given the Zariski topology. Or we think in terms of $\mathbb{C}$-points: a $\mathbb{C}$-point of X is a homomorphism of $\mathbb{C}$-algebras

$$\phi : R \to \mathbb{C}.$$

Notice that such a homomorphism by definition $\phi(1) = 1$, and ϕ is completely determined once we specify the values $\phi(x_i) = a_i \in \mathbb{C}$ for $i = 1, \ldots n$. Thus a $\mathbb{C}$-point determines a tuple

$$(a_1, \ldots, a_n) \in \mathbb{C},$$

and $\mathbb{C}$-points of X are identified with the tuples such that

$$f(a_1, \ldots, a_n) = 0$$

for all $f \in I$. Thus the set of $\mathbb{C}$-points of X is identified with the so-called zero set $V = V(I) \subset \mathbb{C}^n$ of the ideal.

Hilbert's Nullstellensatz says that V is nonempty[4], and furthermore that the ideal I can almost be completely reconstructed from the set V. The actual statement is that the radical ideal $\sqrt{I}$ can be recovered from V as

$$\{f \in \mathbb{C}[x_1, \ldots, x_n] \mid f(V) = 0\}.$$

Loosely speaking, the term affine algebraic variety will mean an algebraic set $V = V(I) \subset \mathbb{C}^n$, corresponding to some ideal $I \subset \mathbb{C}[x_1, \ldots, x_n]$. Given the above remarks, it is equivalent to consider finitely generated rings R containing no nilpotent elements. We then write $\mathrm{Spec}(R)$ for the corresponding variety.

Just to be explicit:

---

[4]Considering the example of $\mathbb{R}[x]/(x^2 + 1)$ shows that algebraic closure of $\mathbb{C}$ is needed for the statement to hold

Definition 13.1. An algebraic set is by definition a subset of $\mathbb{C}^n$ of the form

$$V(I) = \{p \in \mathbb{C}^n \mid f(p) = 0 \text{ for all } f \in I\}$$

for some subset $I \subset \mathbb{C}[x_1, \ldots, x_n]$ (which might as well be an ideal). An affine algebraic variety is an algebraic set. The Zariski topology on $\mathbb{C}^n$ is the topology whose closed sets are the algebraic sets. The Zariski topology of an algebraic variety is the subspace topology induced from $\mathbb{C}^n$. A constructible set in an affine algebraic variety is a finite union of intersections of closed and open sets.

Let $X = \mathrm{Spec}(R)$ and $Y = \mathrm{Spec}(S)$. A morphism $f : X \to Y$ of affine algebraic varieties consists of a homomorphism $S \to R$, simple as that.

Lemma 13.1. *The image of a constructible set under a morphism of affine algebraic varieties is a constructible set. A constructible set S contains a subset which is open and dense in the closure $\overline{S}$.*

Definition 13.2. A linear algebraic group is an affine algebraic variety G together with morphisms $\mu : G \times G \to G$ (multiplication) and $i : G \to G$ (inverse) which make G (as a set) into a group.

Here the product $G \times G$ appears. In general $X \times Y$ means the algebraic variety associated with the tensor product ring $R \otimes_{\mathbb{C}} S$. (Exercise: convince yourself that the tensor product ring does indeed correspond to the Cartesian product of sets.)

The usual examples of Lie groups give examples of algebraic groups. We have $GL_n(\mathbb{C}) \subset \mathbb{C}^{n^2}$ and $SL_n(\mathbb{C}) \subset \mathbb{C}^{n^2}$. Actually, it is plain to see that $SL_n(\mathbb{C})$ is an algebraic variety, since it is the zero set of the ideal generated by the polynomial

$$\det(A) - 1.$$

But why is $GL_n(\mathbb{C})$ an algebraic variety, being as it is characterised by the open condition $\det(A) \neq 0$? Well there is a trick: we actually define $GL_n(\mathbb{C})$ as the algebraic set

$$\{(A, B) \in \mathbb{C}^{n^2} \times \mathbb{C}^{n^2} \mid AB = I_n\}.$$

The $n = 1$ case of this is just

$$\mathbb{C}^\times = \{(a, b) \in \mathbb{C} \mid ab = 1\} = \mathrm{Spec}(\mathbb{C}[x, y]/(xy - 1)).$$

The relation $R \leftrightarrow \mathrm{Spec}(R)$ is a contravariant functor. Or rather one might say that we have defined the category of affine algebraic varieties to be $\mathcal{C}^{op}$, where $\mathcal{C}$ is the category of finitely generated commutative rings without nilpotents. From this point of view an affine scheme is just an object of the opposite of the category of all commutative rings.

Anyway, the fact that $\mu : G \times G \to G$ is associative can be reformulated purely in terms of the corresponding map

$$\Delta : R \to R \otimes_{\mathbb{C}} R,$$

similarly for $S : R \to R$ corresponding to the inversion in G. The structure that arises is that of a *Hopf algebra*. A Hopf algebra is an algebra together with a comultiplication Δ as above, which should be coassociative, as well as being compatible with the usual multiplication in some way, and there should be an antipode S which should be compatible with the multiplication and comultiplication.

IN CLASS, HERE, I GAVE A LONG MEANDERING INTRODUCTION TO THE NOTION OF HOPF ALGEBRA FROM THE POINT OF VIEW OF WANTING TO BE ABLE TO TENSORISE AND DUALISE LEFT-MODULES OVER AN ASSOCIATIVE ALGEBRA. THIS MOTIVATES THE HOPF ALGEBRA STRUCTURE ON $U(\mathfrak{g})$ TOO.

Let us spell this out in the simple example of $\mathbb{C}^{\times}$. Here $R = \mathbb{C}[x, y]/(xy - 1)$. A $\mathbb{C}$-point of R is a homomorphism $\phi : R \to \mathbb{C}$, which is determined by $\phi(x) = a$. Since $\phi(y) \in \mathbb{C}$ is inverse to a, necessarily $a \neq 0$. The product of the $\mathbb{C}$-points ϕ_a and ϕ_b is by definition the $\mathbb{C}$-point ϕ_{ab}. So now we have to ponder what homomorphism

$$\Delta : R \to R \otimes R$$

satisfies

$$(\phi_a \otimes \phi_b) \circ \Delta = \phi_{ab},$$

and it's pretty clear that $\Delta(x) = x \otimes x$ (and $\Delta(y) = y \otimes y$) does the trick, in fact is the only choice that works.

Now let's look at $SL_2(\mathbb{C})$. The ring of functions is

$$\mathbb{C}[x_{11}, x_{12}, x_{21}, x_{22}]/(\det)$$

where $\det = x_{11}x_{22} - x_{12}x_{21}$. By thinking about the structure of the product

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}$$

it becomes clear that

$$\Delta(x) = x_{11} \otimes x_{11} + x_{12} \otimes x_{21},$$

etc. Anyway, the Hopf algebra perspective is interesting, but we won't really use it for anything in this course.

Above we identified the affine algebraic variety as the set of $\mathbb{C}$-points of $X = \mathrm{Spec}(R)$, namely the set of homomorphisms $R \to \mathbb{C}$. This perspective can be enlarged in a natural way: For another $\mathbb{C}$-algebra S one defines an S-point of X to be a homomorphism $R \to S$. Consider what happens if we take $S = \mathbb{C}[\varepsilon]/(\varepsilon^2)$. Then a homomorphism $\varphi : R \to S$ is of the form

$$\phi(r) = \phi_0(r) + \varepsilon\phi_1(r),$$

where $\phi_0, \varphi_1 : R \to \mathbb{C}$. It is easy to see that ϕ_0 itself has to be a homomorphism (so is nothing other than a $\mathbb{C}$-point of X), but $\varphi_1 : R \to \mathbb{C}$ is $\mathbb{C}$-linear but not a homomorphism. Indeed

$$\phi_1(rs) = \phi_0(r)\phi_1(s) + \phi_1(r)\phi_0(s).$$

Something that is clear from this is that $\phi_1(1) = 0$. Let us consider a point $e \in X$ and consider only S-points whose "underlying $\mathbb{C}$-point" is e. Then it is clear that ϕ_1 is completely determined by its values on $I \subset R$ the ideal of functions which vanish at e. On the other hand, ϕ_1 descends to a well-defined and linear function on the quotient

$$I/I^2.$$

Those who know a little algebraic geometry will recognise that I/I^2 is the Zariski cotangent space, and so ϕ_1 amounts to precisely an element of the Zariski tangent space $T_eX = (I/I^2)^*$.

Now suppose that $X = G$ is a linear algebraic group. The Hopf algebra structure on R should translate into some extra structure on T_eG, which (you might have guessed) ends up being the structure of Lie algebra. Let's leave this as an exercise.

A morphism of algebraic groups induces a morphism of the corresponding Lie algebras, and many concepts in the two worlds match up: Lie subalgebras with subgroups, ideals with normal subgroups. The quotient of a Lie algebra by an ideal has a natural Lie algebra structure. One might ask the same about algebraic groups. It turns out that yes, the quotient of a linear algebraic group by a closed normal subgroup carries a structure of linear algebraic group. We will prove it later.

In practice this perspective gives a very direct construction of the corresponding Lie algebra. For example the Lie algebra of $SL_n(\mathbb{C})$ should be identified with the set of matrices $X \in Mat_n(\mathbb{C})$ such that

$$I_n + \varepsilon X \in SL_n(S).$$

But one can check that $\det(I_n + \varepsilon X) = 1 + \varepsilon \operatorname{Tr}(X)$, so we recover

$$\mathfrak{sl}_n(\mathbb{C}) = \{X \in Mat_n(\mathbb{C}) \mid \operatorname{Tr}(X) = 0\}.$$

Or for example

$$SO_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) \mid A^T A = I_n\},$$

has Lie algebra consisting of X for which

$$(I_n + \varepsilon X^T)(I_n + \varepsilon X) = I_n,$$

i.e., $X + X^T = 0$.

Remark 13.2. *Every closed subgroup of a linear algebraic group induces a Lie subalgebra. But not every Lie subalgebra corresponds to a closed subgroup. The simplest example occurs for $G = (\mathbb{C}^\times)^2$ whose Lie algebra is $\mathbb{C}^2$ with Lie bracket 0. Any vector subspace is a Lie subalgebra but a vector subspace $(t, \alpha t)$ (for some fixed irrational number α) can only correspond to a subgroup of the form $\{(z, z^\alpha)\}$. Bu such a set is not algebraic.*

Now let us recall some basic notions of projective geometry. The n-dimensional projective space $\mathbb{P}^n$ is by definition the set

$$\mathbb{P} = (\mathbb{C}^{n+1} \setminus (0, \ldots, 0)) / \sim,$$

where the equivalence relation $\sim$ is defined by $v \sim \lambda v$ for $\lambda \neq 0$. We think of $\mathbb{P}^n$ as the set of lines through the origin in $\mathbb{C}^{n+1}$. We tend to write points of $\mathbb{P}^n$ in the format $[v_0 : v_1 : \ldots : v_n]$. We write

$$D_i = \{[v_0 : v_1 : \ldots : v_r] \mid v_i \neq 0\},$$

and we observe that there is an isomorphism $D_i \to \mathbb{C}^r$ via

$$[v_0 : v_1 : \ldots : v_r] \mapsto (\frac{v_0}{v_i}, \frac{v_1}{v_i}, \ldots, \frac{\widehat{v_i}}{v_i}, \ldots, \frac{v_r}{v_i}).$$

Inspired by the definition of smooth manifold, a projective variety may be defined as a subset of $\mathbb{P}^n$ whose intersection with each open affine U_i is an affine algebraic variety. A good starting point to get intuition about projective varieties is that they are analogous to compact topological spaces (affine algebraic varieties being never compact).

Definition 13.3. Let X be a complex algebraic variety. We say that X is proper (or complete) if it is compact and Hausdorff in the complex analytic topology.

For example a projective variety is proper. Hironaka gave a (difficult) example of a 3-dimensional smooth and proper but not projective variety. Some facts about proper varieties:

- A closed subvariety of a proper variety is proper.

- An affine and proper variety is zero dimensional (i.e., consists of a finite collection of points).

We are only discussing linear algebraic groups, which by definition are affine. But when we come to consider quotient varieties, projective varieties arise naturally.

Give examples of projective varieties: explain why Grassmannians and flag varieties are projective. Via the Plucker embedding, Grassmannians are projective varieties.

Maybe here is a good place to talk about sheaves of $\mathcal{O}$-modules, in aid of the definition of scheme in general?

Definition 13.4. An action of an algebraic group G on a variety X is a morphism of algebraic varieties $G \times X \to X$ (written $(g, x) \mapsto gx$) such that

- $g_1(g_2 x) = (g_1 g_2) x$

- $ex = x$.

This lemma is [**?**, 8.2].

Lemma 13.3. *Let G act on X. The set of fixed points in X of any $g \in G$ is a closed subvariety. The same is true for the fixed points of G.*

Proof. Let $g \in G$ and consider the morphism $\psi : X \to X \times X$ given by $x \mapsto (x, gx)$. The set of fixed points of g is precisely the inverse image of the diagonal $\Delta \subset X \times X$ under ψ. The diagonal is closed (in the Zariski topology) and the preimage of a closed set is closed.

The fixed points of G is just the intersection of all the sets of fixed points, and arbitrary intersections of closed sets are closed. $\square$

This lemma is [**?**, 8.3].

Lemma 13.4. *Let G act on a variety X. Then each orbit is smooth and its boundary is a union of orbits of strictly lower dimension. In particular there exists a closed orbit in X.*

Proof. We first observe that the closure $\overline{Y}$ is invariant under G. Indeed $\overline{Y}$ is the union of all closed sets containing Y, but the collection of closed sets containing Y is invariant under the action of G and hence so is their mutual intersection. It follows immediately that $\overline{Y} \backslash Y$ is a union of orbits.

Let $Y = Gy$ for some $y \in X$. We now wish to show that Y is open in $\overline{Y}$. Since it is the image of G under the map $g \mapsto gy$, the set Y is constructible, and so contains a subset S which is open and dense in $\overline{Y}$. The union of all translates by G of S is a union of sets open in $\overline{Y}$, hence open in $\overline{Y}$ and at the same time this union equals Y. Thus Y is open in $\overline{Y}$.

Hence the complement $\overline{Y} \backslash Y$ is closed, hence (assuming we know a thing or two about dimension) of strictly lower dimension. Since the dimension cannot go down forever, there must be some orbit for which $\overline{Y} \backslash Y$ is empty, i.e., a closed orbit. $\square$

An algebraic group G is said to be solvable if it is solvable as an abstract group, i.e., one defines (G, G) to be the subgroup of G generated by $ghg^{-1}h^{-1}$ for $g, h \in G$, and G is solvable if the derived series $D^0(G) = G$, $D^1(G) = (G, G)$, $D^{n+1}(G) = (D^n(G), D^n(G))$, eventually goes to $\{e\}$.

The derived series corresponds at the level of Lie algebras to the derived series in that context, and

Definition 13.5. Let G be a linear algebraic group. A Borel subgroup $B \subset G$ is a closed connected solvable subgroup which is maximal with respect to inclusion.

Definition 13.6. Let $\mathfrak{g}$ be a Lie algebra. A Borel subalgebra $\mathfrak{b} \subset \mathfrak{g}$ is a maximal solvable Lie subalgebra.

In the following theorem, it is important that the ground field (which for us is always $\mathbb{C}$ anyway) be algebraically closed and of characteristic 0.

Theorem 13.5 (Lie's theorem)**.** *Let $\mathfrak{g}$ be a solvable Lie subalgebra of gl_V. Then there exists a one-dimensional $\mathfrak{g}$-submodule $L \subset V$.*

For the following theorem, see [**?**, 17.6]. Apparently, and weirdly, it is true even if the base field is not of characteristic zero.

Theorem 13.6 (Lie-Kolchin theorem). *Let G be a connected solvable subgroup of $GL(V)$. Then there exists a one-dimensional G-submodule $L \subset V$.*

This proposition is [**?**, 21.2].

Proposition 13.7. *Let G be a connected algebraic group, and let X be a nonempty proper variety with an action of G. If G is solvable then the action has a fixed point.*

Proof. The proof is by induction on the dimension of G. If $\dim G = 0$ then since it is connected $G = \{e\}$ and so X is all fixed points.

Now suppose $\dim G > 0$ and let $H \subset G$ be the commutator subgroup, which has strictly smaller dimension because of solvability. By the inductive hypothesis, H has a fixed point on X. So let us set $Y \subset X$ the set of H-fixed points.

We know that Y is closed by Lemma **??**. In general a closed subvariety of a proper variety is proper. Since $H \subset G$ is a normal subgroup, the action of G on X preserves Y. So we have a proper variety Y with an action of G, in which every point is fixed by H.

For any $x \in Y$ the stabiliser $\mathrm{Stab}_G(x)$ contains H and so is a normal subgroup of G. According to Theorem **??**, this implies that $G/Stab_G(x)$ is an affine variety. By Lemma **??** one can choose Gx closed (hence proper). So we have a bijection

$$G/\mathrm{Stab}_G(x) \to Gx$$

from an affine variety to a proper variety. But an affine proper variety is a point, so $G = \mathrm{Stab}_G(x)$[5] $\qquad \square$

By applying the theorem to the projectivisation $\mathbb{P}(V)$, the Lie-Kolchin theorem follows immediately.

Now we come to the main theorem.

Proposition 13.8. *Let G be a linear algebraic group, and $B \subset G$ a Borel subgroup. Then the quotient space G/B carries the structure of a smooth projective variety, compatible with the action of G.*

First some preliminaries. This one is [**?**, 11.2]

Proposition 13.9. *Let G be a linear algebraic group, and $H \subset G$ a closed subgroup. There exists a rational representation $\rho : G \to GL(V)$ and a one-dimensional subspace $L \subset V$ such that H is exactly the stabiliser of L, i.e.,*

$$\{g \in G \mid \rho(g)L = L\} = H.$$

---

[5] Actually, in [**?**] it is explained that this is not entirely trivial, so a careful proof is given there.

Proof. Let $I \subset \mathbb{C}[G]$ denote the vanishing ideal of H. As for any ideal in a Noetherian ring, I is finitely generated, so [Humphreys Lin Alg Grps, Prop 8.6] we can find a finite dimensional G-invariant subspace $W \subset \mathbb{C}[G]$ which contains a set of generators of I.

Basic idea: actually more generally for an algebraic action of G on affine variety V, it amounts to a morphism $\mathbb{C}[V] \to \mathbb{C}[V] \otimes \mathbb{C}[G]$, and the image of any $f \in \mathbb{C}[V]$ is a finite linear combination of $f_i \otimes g_i$. The subspace spanned by the f_i contains all G-translates of f. Nice.

Now set $M = W \cap I$.

By [Humphreys Lin Alg Grps, Prop 8.5] apparently $\{g \in G \mid g(I) = I\} = H$. One direction is obvious. For the other direction, if $g(I) = I$ and $f \in I$ then by definition $f(x) = 0$ for all $x \in H$ including $x = e$, now $[g.f](x) = f(g^{-1}x)$ so $g.f$ vanishes at g^{-1}. So the hypothesis $g.f \in I$ implies $g \in H$ in fact.

Therefore $H(M) = M$. Now we claim that $\{g \in G \mid g(W) = W\} = H$ in fact. This is easy since $I = M\mathbb{C}[G]$, so if $g(M) = M$ we deduce $g(I) = g(M\mathbb{C}[G]) = g(M)\mathbb{C}[G] = I$, so $g \in H$.

Let $d = \dim(M)$ and set $V = \wedge^d(W)$ and $L = \wedge^d(M)$. A simple lemma asserts that $\wedge^d \rho(g)$ fixes L if and only if $\rho(g)$ fixes M, so we are done. $\qquad \square$

As a direct corollary we have:

Proposition 13.10. *The quotient of an affine algebraic group by a closed subgroup, is an algebraic variety (quasiprojective).*

Proof. We let V and L be as in the previous proposition. We can think about the action of G on $\mathbb{P}V$, and the orbit $G \cdot [L]$. The stabiliser of $[L]$ is precisely H, so we have $G \cdot [L] = G/H$ as sets. Thus we have presented G/H as the image of a morphism of algebraic varieties, so it is a constructible subset of projective space. $\qquad \square$

Another important corollary is the following (see [**?**, Proposition 5.5.10]).

Theorem 13.11. *The quotient G/H of a linear algebraic group G by a closed normal subgroup H is again a linear algebraic group.*

Proof. By Proposition **??** there is a representation (V, ρ) of G and a line $L \subset V$ for which we have $N = Stab_G(L)$. The action of H on L then constitutes a character, i.e.,

$$\rho(h)v = \chi_0(h)v$$

for some $\chi_0 : H \to \mathbb{C}^\times$. Define, for each character χ, the subspace V_χ consisting of vectors v such that $\rho(h)v = \chi(h)v$ for all $h \in H$. So $L \subset V_{\chi_0}$ in particular. We now take $V' \subset V$ to be the span of all these V_χ. Since H is normal in G, each $\rho(g)$ permutes the V_χ and hence V' is a subrepresentation of V. Assume without loss of generality that $V' = V$.

Now define $W \subset \text{End}(V)$ by

$$W = \{f : V \to V \mid f(V_\chi) \subset V_\chi \text{ for each } \chi\}.$$

We consider the conjugation representation of G in W, namely

$$\psi(g)f = \rho(g) \circ f \circ \rho(g)^{-1}.$$

Since each $\rho(h)$ acts by some constant in each V_χ, we have $\psi(h) = I_W$ for all $h \in H$. Let us now consider $g \in \ker(\psi)$. Firstly it is clear that such g must preserve each V_χ. But then by Schur's lemma, its restriction to each V_χ must be by some constant. In particular $\rho(g)$ fixes L and so $g \in H$.

Thus we have an injective map $G/H \to GL(W)$ which, since it is a representation, is a homomorphism of groups.

Springer makes a point of checking the differential too, which is a map of Lie algebras. Not sure if this is just a characteristic p thing or if it is necessary too in the complex case. $\quad\square$

of Theorem **??**. Let B be a Borel in G of largest possible dimension. By Proposition **??** there exists a G-module V and a line $L \subset V$ for which B is the stabiliser of L.

We now look at the quotient module V/L of B. By Theorem **??** there is another invariant line L_2 here, and by continuing in this way we construct a flag

$$L = V_1 \subset V_2 \subset \ldots \subset V$$

fixed by B. In fact by construction of L, the stabiliser of this flag (call it F) is exactly B.

Thus we have identified G/B with the orbit $G \cdot F$ in the flag variety $Fl(V)$. The stabiliser of any point in $Fl(V)$ is a solvable subgroup of G (WHY? Presumably because intersection of G with upper triangular matrices from $GL(V)$), hence of dimension no greater than that of B. It follows that $G \cdot F$ has minimal dimension, and hence is closed by Proposition **??**. As the flag variety is projective, so is G/B. $\quad\square$

OK so why is $Fl(V)$ projective? This is easy,

Firstly consider the Grassmannian $Gr(n, d)$ of d-planes inside $\mathbb{C}^n$. We associate to such a plane the form $v_1 \wedge \ldots \wedge v_d$, which determines the d-plane and is in turn only well-defined up to scalars. Thus there is an embedding

$$Gr(n, d) \to \mathbb{P}(\wedge^d \mathbb{C}^n).$$

The flag variety consists of a series of planes of dimension 1, 2, etc. related by the conditions of inclusion. So $Fl(V)$ is a closed subvariety of

$$Gr(n, 1) \times Gr(n, 2) \times \cdots \times Gr(n, n-1).$$

14 From integral weights to line bundles on G/B

We are now going to describe a procedure which takes an integral weight $\Lambda \in P \subset \mathfrak{h}^*$ and produces a line bundle over G/B. A very efficient reference for this is [?, p. 253].

Let $T \subset B \subset G$ be a maximal torus and Borel subgroup of a semisimple algebraic group $G \subset GL_n(\mathbb{C})$. For intuition, one may imagine $G = SL_n(\mathbb{C})$ with T the diagonal matrices and B the upper triangular.

We know the one-dimensional representations of $\mathfrak{h} = \mathbf{Lie}(T)$ are just given by weights $\lambda \in \mathfrak{h}^*$, i.e., $\mathbb{C}v_\lambda$ with

$$Hv_\lambda = \lambda(H)v_\lambda.$$

Intuitively we would like to "exponentiate" such a representation to get a representation of T. Informally, an element of T can be written e^H for some $H \in \mathfrak{h}$, and we would like to define the action of e^H on v_λ as being

$$e^H v_\lambda = e^{\lambda(H)} v_\lambda.$$

This will not work for an arbitrary λ, indeed consider SL_3 and sl_3: the matrices

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

we have $e^{2\pi i H_1} = e^{2\pi i H_2} = I_3$, and so we had better have $\lambda(H_1), \lambda(H_2) \in 2\pi i \mathbb{Z}$, in other words we need $\lambda \in \mathbb{Z}\varpi_1 + \mathbb{Z}\varpi_2$.

This is more-or-less the general situation: only weights in some lattice $L \subset \mathfrak{h}^*$ will exponentiate to representations of the torus T, in this case the lattice was P, but in general there are a small finite number of possibilities $Q \subset L \subset P$.

More concretely now, for a weight $\lambda = m\varpi_1 + n\varpi_2$, a general element of $T \subset SL_3$ looks like

$$h = \begin{pmatrix} s & 0 & 0 \\ 0 & s^{-1}t & 0 \\ 0 & 0 & t^{-1} \end{pmatrix} \quad \text{and} \quad hv_\lambda = s^m t^n v_\lambda.$$

So let Λ be such an integral weight. We extend $M = \mathbb{C}v_\Lambda$ module over the Borel subgroup $B \subset G$ by letting $u \in U$ act as the identity. This is the exponentiated version of the one-dimensional $\mathfrak{b}$-module defined by $xv = \Lambda(x)v$ for $x \in \mathfrak{h}$ and $xv = 0$ for $x \in \mathfrak{n}_+$. In fact for the construction below we could take any B-module M.

Now we construct a vector bundle over $X = G/B$ as follows:

$$E = E_\Lambda = (G \times M)/\sim, \qquad (gb, v) \sim (g, bv).$$

The projection map $p : E \to X$ is given by $(g, v) \mapsto gB$. Notice that G acts on the total space E in a manner compatible with the projection p, namely $g_1(g, v) = (g_1 g, v)$.

The fibre $p^{-1}(gB)$ consists of the image in E of all pairs $(gb, v) \in G \times M$ for $b \in B$. All of these are identified in E with pairs of the form (g, v). On the other hand all such pairs occur and are mutually distinct, so $p^{-1}(gB)$ is in bijection with M.

If we had chosen a different representative $g' \in gB$ (so $g' = gb_0$ for some $b_0 \neq e$) then the bijection $p^{-1}(gB) \to M$ obtained would be different.

How best to show this is locally trivial? It suffices to do so near $g = e$. We should use the fact that $U^- B$ is dense in G. Maybe the best way to show this would be to observe that the differential
$$T_e U^- \otimes T_e B \to T_e G$$
is surjective.

A section of E is a function $s : X \to E$ such that $p \circ s = id_X$. We can always write $s(gB)$ in the form $(g, \varphi(g))$, and we are led to consider the function

$$\varphi : G \to M.$$

In fact we can be even more concrete, we write $\varphi(g) = f(g)v_\Lambda$. Evidently $s(gbB) = s(gB)$, but our translation to functions has given us

$$(g, f(g)v_\Lambda) = s(gB) = s(gbB) = (gb, f(gb)v_\Lambda) = (g, f(gb)\Lambda(b)v_\Lambda),$$

so $f(gb) = \Lambda(b)^{-1} f(g)$. Thus $\Gamma(G/B, \mathcal{L}_\Lambda)$ is identified with

$$\{f : G \to \mathbb{C} \mid f(gb) = \Lambda(b)^{-1} f(g) \text{ for all } b \in B\},$$

There is a natural action of G on this set of functions, in which $g_1 f$ is defined by

$$[g_1 f](g) = f(g_1^{-1} g).$$

15   The Borel-Weil theorem

Theorem 15.1 (Borel-Weil)**.** *The space of global sections $\Gamma(G/B, \mathcal{L}_\Lambda)$ is always a $\mathfrak{g}$-module in a natural way. If $\Lambda \in P_+$ then*

$$\Gamma(G/B, \mathcal{L}_{-\Lambda}) \cong L(\Lambda)^*,$$

and for all other Λ is is zero.

The note of Lurie [**?**] provides a useful sketch of the proof.

Before presenting the proof, let's work out the case of $G = SL_2$ in detail. Let's warm up by checking that

$$\mathcal{L}_{-2\varpi} \cong T_{\mathbb{P}^1}.$$

77

In the big cell we have the coordinate u as above and the fibres are of the form

$$\left(\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}, \mathbb{C}v_\lambda\right).$$

As we have computed in a previous lecture

$$\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} u^{-1} & 1 \\ 0 & u \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ u^{-1} & 1 \end{pmatrix}.$$

We trivialise the bundle in a neighbourhood of $\{\infty\}$ as

$$\left(\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}, \mathbb{C}v_\lambda\right).$$

On the overlap we have $x = 1/u$ and

$$\left(\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}, v_\lambda\right) = \left(\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}, u^{-k}v_\lambda\right)$$

If we take $k = -2$, then this is just the transformation law for $T_{\mathbb{P}^1}$ where

$$\left(\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}, v_\lambda\right) \equiv -\frac{d}{dx} \qquad \left(\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}, v_\lambda\right) \equiv \frac{d}{du}.$$

Now let's think about the action of $\mathfrak{g} = sl_2$ via vector fields on $\mathbb{P}^1$ more generally. Write $X = G/B \cong \mathbb{P}^1$ for the flag variety (or any variety at this point), and consider a smooth function $f : X \to \mathbb{C}$. We have an action of $g \in G$ by

$$[gf](x) = f(g^{-1}x),$$

and we can differentiate to obtain an action of $\mathfrak{g}$, i.e.,

$$[af](x) = \frac{d}{dt}f(e^{-ta}x)|_{t=0}.$$

In the present case, to be concrete, let us restrict functions to the big cell, which is identified with $\mathbb{C}$ by

$$u \leftrightarrow \begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix} B/B.$$

The action of E is as follows

$$\frac{d}{dt}f(e^{-tE}x)|_{t=0} = \frac{d}{dt}f\left(\begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}\right)|_{t=0}$$

$$= \frac{d}{dt}f\left(\begin{pmatrix} u-t & -1 \\ 1 & 0 \end{pmatrix}\right)|_{t=0} = \frac{d}{dt}f(u-t)|_{t=0}.$$

78

Thus
$$E \mapsto -\frac{d}{du}.$$

The cases of H and F are a little trickier:

$$\frac{d}{dt}f(e^{-tF}x)|_{t=0} = \frac{d}{dt}f\left(\begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}\right)|_{t=0}$$
$$= \frac{d}{dt}f\left(\begin{pmatrix} u & -1 \\ 1-ut & t \end{pmatrix}\right)|_{t=0}.$$

To proceed we have to write the product matrix in the standard form by multiplying on the right with an element of B. In fact

$$\begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} u(1-tu)^{-1} & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1-tu & t \\ 0 & (1-tu)^{-1} \end{pmatrix},$$

so F acts as

$$\frac{d}{dt}f(u(1-tu)^{-1})|_{t=0},$$

or rather

$$F \mapsto u^2\frac{d}{du}.$$

A similar calculation yields

$$H \mapsto -2u\frac{d}{du}.$$

Now let us examine the action of these operators on the sections of one of the line bundles $\mathcal{L}_\lambda$. Over the big cell a section of the line bundle can be written as

$$u \mapsto s(u) = f(u)(\begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}, v_\lambda).$$

We now repeating the calculation for F noting that now

$$(\begin{pmatrix} u(1-tu)^{-1} & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1-tu & t \\ 0 & (1-tu)^{-1} \end{pmatrix}v_\lambda) = (\begin{pmatrix} u(1-tu)^{-1} & -1 \\ 1 & 0 \end{pmatrix}, (1-tu)^k v_\lambda)$$

so in terms of the function $f(u)$ we have

$$\frac{f(u(1-tu)^{-1})}{(1-tu)^k}$$

(The $(1-tu)^k$ turning into $(1-tu)^{-k}$ is a tricky point.) So F acts as

$$u^2\frac{d}{du} + ku.$$

In full we have

$$E \mapsto -\frac{d}{du},$$

$$H \mapsto -2u\frac{d}{du} - k,$$

$$F \mapsto u^2\frac{d}{du} + ku.$$

For $k = -N$, DIAGRAM OF CHAIN $1, u, u^2, \ldots, u^N$ WITH OPERATORS E, H, F.

The proof of the Borel-Weil theorem requires the following important result.

Theorem 15.2. *Let X be a smooth projective algebraic variety (or a compact complex manifold) and E a holomorphic vector bundle over X. Then the vector space of holomorphic global sections $\Gamma(X, E)$ is finite dimensional.*

Proof of the Borel-Weil theorem. We have established that G/B is a projective algebraic variety, and that $\mathcal{L}_\Lambda$ is a holomorphic line bundle. Therefore the vector space $\Gamma(X, \mathcal{L}_\Lambda)$ is finite dimensional by Theorem **??**, recall that it is identified with

$$V = \{f : G \to \mathbb{C} \mid f(gb) = \Lambda(b)^{-1}f(g) \text{ for all } b \in B\},$$

with the action of G being

$$[gf](x) = f(g^{-1}x).$$

By differentiation this space becomes a $\mathfrak{g}$-module. The first observation is that since V is finite dimensional, we have a constraint on $\Lambda \in P$.

Indeed let

$$V^- = \{f \in V \mid f(ux) = f(x) \text{ for all } u \in U^-\}.$$

For any $f \in V^-$ we have

$$f(ub) = f(1b) = \Lambda(b)^{-1}f(1).$$

If we had two functions $f_1, f_2 \in V^-$ with $f_2(1) \neq 0$ then we would have

$$f_1(x) = \frac{f_1(1)}{f_2(1)}f_2(x)$$

for all $x \in U^-B$. But

$$U^-B = U(w_\circ)n_{w_\circ}B = Bw_\circ B \subset G$$

is the big double coset. Since this double coset is dense in G, the proportionality relation between f_1 and f_2 above, must hold for all $x \in G$. So $\dim(V^-) \leq 1$.

We can actually conclude that V is irreducible already. If it were not then it would have strictly more than one singular vector for the action of $\mathfrak{n}_-$. But such singular vectors are just the same as elements of V^-, which we have just shown to be at most 1-dimensional.

80

Further we claim that the weight of $f \in V^-$ is Λ, and so the only way that $V \neq 0$ might hold is if $-\Lambda \in P_+$. Indeed for $k \in T$ we compute (for arbitrary $u \in U^-$ and $b \in B$)

$$[kf](ub) = f(k^{-1}ub) = f(u'k^{-1}b) = \Lambda(k)f(u'b) = \Lambda(k)f(ub).$$

(Here we used that U^- is normalised by T, as is B, and we used $f(u'b) = f(ub) = f(b)$).

It remains to prove $V^- \neq 0$ when $\Lambda \in P_+$. To do this we shall construct a nonzero holomorphic function $f : G \to \mathbb{C}$ such that $f(ub) = \Lambda(b)^{-1}$ for all $u \in U^-$ and $b \in B$. It is easy to see that such a function satisfies the required B-invariance condition, and so will be an element of V (in fact also V^-).

Such a function f (if it exists) is already determined on all of $U^-B \subset G$, with

$$f(ub) = \Lambda(b)^{-1}.$$

We recall the decomposition

$$G = \bigcup_{w \in W} BwB \tag{15.1}$$

The "big double coset" is $U^- \times B$. We wish to show that f is regular on double cosets of lower dimension. In fact, by a result called Hartog's theorem, it suffices to do this for double cosets of codimension 1. These correspond to the simple roots of $\mathfrak{g}$.

Let us sketch the argument for $G = SL_n$. To be concrete let's suppose the character corresponds to the weight $\Lambda = \sum_i k_i \varpi_i$ where all $k_i \in \mathbb{Z}$.

So let α be a simple root, and let $P \subset G$ be the corresponding parabolic subgroup, namely the subgroup generated by U^- and $\exp(tE_\alpha)$. In $G = SL_n(\mathbb{C})$, for example, P consists of matrices that look something like the following:

$$\left(\begin{array}{c|cc|cc} 1 & 0 & 0 & 0 & 0 \\ \hline * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ \hline * & * & * & 1 & 0 \\ * & * & * & * & 1 \end{array} \right)$$

Let $G_0 \cong SL_2(\mathbb{C})$ denote the block diagonal matrices with two-by-two block as indicated. We now consider the restriction of our function f to $G_0 \subset G$. Any element

$$\left(\begin{array}{cc} a & b \\ c & d \end{array} \right) \in G_0$$

in which $a \neq 0$ can be written in the form

$$\left(\begin{array}{cc} 1 & 0 \\ c & 1 \end{array} \right) \left(\begin{array}{cc} a & b \\ 0 & 1/a \end{array} \right) = \left(\begin{array}{cc} a & b \\ ac & bc + 1/a \end{array} \right),$$

81

and the value of f on any such element is a^{-k_i}.

We go into the lower dimensional double coset by taking the limit

$$\begin{pmatrix} 1 & 0 \\ -1/(b_0 a) & 1 \end{pmatrix} \begin{pmatrix} a & b_0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} a & b_0 \\ -1/b_0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & b_0 \\ -1/b_0 & 0 \end{pmatrix}$$

as $a \to 0$. The assumption $-\Lambda \in P_+$ now says that the value of f vanishes in this limit, as required. $\qquad\square$

16   Finiteness of global sections: analytic proof

Here is an elementary proof of Theorem **??**, in case the projective variety X is 1-dimensional. Such a variety is the same thing as a compact Riemann surface. Let us use the notation $\Delta_r = \{z \in \mathbb{C} \mid |z| < r\}$ for positive real r, and $\Delta = \Delta_1$.

Proposition 16.1 (Maximum modulus principle). *Let $U \subset \mathbb{C}$ be some open set containing the closure $\overline{\Delta}_r$, and $f : U \to \mathbb{C}$ holomorphic. Then there exists $z_0 \in \partial \Delta_r$ such that*

$$\sup_{z \in \Delta_r} |f(z)| \leq |f(z_0)|.$$

The Schwarz lemma is a useful corollary of the maximum modulus principle.

Proposition 16.2 (Schwarz lemma). *Let $U \supset \overline{\Delta}$ and $f : U \to \mathbb{C}$ holomorphic, and suppose that $f(0) = 0$ and $|f(z)| \leq 1$ for all $z \in \Delta$. Then*

$$|f(z)| \leq |z|$$

for all $z \in \Delta$.

Proof. We consider the holomorphic function $g(z) = f(z)/z$ and apply the maximum modulus principle to g for each $0 < r < 1$. There exists $z_r \in \partial \Delta_r$ such that

$$|g(z)| \leq |g(z_r)| = \frac{|f(z_r)|}{|z_r|} = \frac{|f(z_r)|}{r} \leq \frac{1}{r}$$

for all $z \in \Delta_r$. Taking the limit as $r \to 1$ yields

$$|g(z)| \leq 1$$

for all $z \in \Delta$. But this tells us that $|f(z)| \leq |z|$ for all $z \in \Delta$. $\qquad\square$

The same proof, with $g(z) = f(z)/z^k$ yields the following more general version:

Proposition 16.3. *Let $U \supset \overline{\Delta}$ and $f : U \to \mathbb{C}$ holomorphic, and suppose that $f(z)$ vanishes to order k at $z = 0$, and that $|f(z)| \leq A$ for all $z \in \Delta$. Then*

$$|f(z)| \leq A|z|^k$$

for all $z \in \Delta$.

Now the finiteness theorem. This proof can be found in [**?**, p. 32]. It is only given for Riemann surfaces Same proof works for higher dimensional complex varieties?

Theorem 16.4. *Let X be a compact Riemann surface, and E a holomorphic vector bundle over X. Then the vector space $\Gamma(X, E)$ of global holomorphic sections of E, is finite dimensional.*

Proof. Let us cover X by charts U_i, each homeomorphic to $\Delta \subset \mathbb{C}$ with corresponding coordinate $z_i : U_i \to \Delta$. Let $a_i = z_i^{-1}(0) \in U_i$, and also let $V_i = z_i^{-1}(\Delta_{1/2}) \subset U_i$.

We assume each chart U_i is sufficiently small that its closure is contained in another chart $W_i \subset X$ on which E has a trivialisation

$$\tau_i : \pi^{-1}(W_i) \to W_i \times \mathbb{C}^n.$$

We also assume that the V_i cover X. And since X is compact, we do all of this with a finite set $U_1, \ldots, U_N$ of charts. The trivialisations of E give holomorphic functions

$$g_{ij} = \tau_i \circ \tau_j^{-1} : W_i \cap W_j \to GL_n(\mathbb{C}).$$

Let $s \in \Gamma(X, E)$. Its restrictions to W_i give functions $s_i : W_i \to \mathbb{C}^n$ with the property $s_i(x) = g_{ij}(x)s_j(x)$ on intersections. Now consider the norms

$$||s||^U = \max_i \sup_{U_i} |s_i(x)|$$

and similarly $||s||^V$ using the V. You might think $||s||^U = ||s||^V$ because both the $\{U_i\}$ and the $\{V_i\}$ cover X, but this is not true because the $|s_i(x)|$ make use of the choice of charts. That is, even for a single point $x \in X$ the values of $|s_i(x)|$ and $|s_j(x)|$ might be different.

What is obvious is that $||s||^V \leq ||s||^U$, and what we can further say is that there exists a constant $C > 0$ such that

$$||s||^U \leq C||s||^V.$$

Indeed, because the closure of U_i is compact, $\sup_{U_i} |s_i(x)|$ is realised as $|s_i(x_0)|$ for some $x_0 \in \overline{U}_i$. Then $x_0 \in V_j$ for some j, and

$$|s_i(x_0)| = |g_{ij}(x_0)s_j(x_0)| \leq C|s_j(x_0)|,$$

where $C = \max_{i,j} \sup_{U_i \cap U_j} |g_{ij}(x)|$ which is finite.

Consider the ring $\mathcal{O}_i$ of germs of holomorphic functions at the point $a_i \in X$, with its ideal m_i of functions which vanish at a_i. The quotient $\mathcal{O}_i/m_i^k$ is a vector space of dimension k. By expanding the function s_i as a power series at the point a_i and taking the first k coefficients, we have a linear map

$$\Gamma(X, E) \to (\mathcal{O}_i/m_i^k)^{\oplus n},$$

in fact by expanding at each a_i, we obtain a linear map

$$\Gamma(X, E) \to \bigoplus_i (\mathcal{O}_i/m_i^k)^{\oplus n}.$$

We would like to show that for some sufficiently large k, this map is injective. This will imply the theorem.

So let $s \in \Gamma(X, E)$, let $k \geq 0$, and suppose $s_i(x)$ vanishes to order k at a_i. Now we invoke the Schwarz lemma: it implies that $|s_i(z)| \leq A|z|^k$ on $U_i \cong \Delta$ for some constant A, which implies $|s_i(z)| \leq 2^{-k}A$ on $V_i \cong \Delta_{1/2}$. Therefore $||s||^V \leq 2^{-k}||s||^U$. So now we have

$$||s||^U \leq C||s||^V \leq 2^{-k}C||s||^U.$$

If we take k sufficiently large that $2^{-k}C < 1$ then the only way this can be satisfied is if $s = 0$. This shows injectivity of the linear map above, for some sufficiently large k. $\qquad \square$

17 Second interlude on sheaves: $\mathcal{O}$-modules

If we are doing geometry then each space X should come with a special sheaf $\mathcal{O}_X$ of what we call "functions on X". This should be a sheaf of rings on X. The intuition of functions is that whenever $f : X \to Y$ is a morphism in whatever category of spaces we are studying, there must be a notion of pullback of a function. That is, there ought to be a morphism

$$f^{-1}\mathcal{O}_Y \to \mathcal{O}_X.$$

By adjointness there is a corresponding distinguished morphism

$$\mathcal{O}_Y \to f_*\mathcal{O}_X.$$

We have already seen the fundamental example of X a topological space, and C_X the sheaf of all continuous functions on X. Below we will discuss another fundamental example: the structure sheaf on the spectrum of a ring.

A simple example of pullback now.

Example 1. We consider X the space consisting of a single point, and $Y = \operatorname{Spec}\mathbb{C}[t]$ the complex line, and $i : X \to Y$ the map sending the point to the origin $\{0\}$. Then $\mathcal{O}_X(X) = \mathbb{C}$ and $\mathcal{O}_Y(Y) \cong \mathbb{C}[t]$. The pullback $i^{-1}\mathcal{O}_Y$ can be defined stalkwise, and in fact must just assign the stalk $\mathcal{O}_{Y,0}$ to X. The stalk, by its definition, is a certain colimit. Intuitively it

consists of rational functions $f(t)$ with no pole at $t = 0$. Let us write R_0 for this localisation. The map $i^{-1}\mathcal{O}_Y \to \mathcal{O}_X$ now consists of

$$R_0 \to \mathbb{C}$$

which is just the act of evaluation $f(t) \mapsto f(0)$.

Definition 17.1. Let X be a topological space and $\mathcal{O}_X$ a sheaf of rings on X. An $\mathcal{O}_X$-module is a sheaf $\mathcal{F}$ on X in which $\mathcal{F}(U)$ is an $\mathcal{O}_X(U)$-module, and the restriction maps are compatible with the module structures.

Let $\mathcal{F}$ be a sheaf of $\mathcal{O}_X$-modules. The sheaf pushforward $f_*\mathcal{F}$ is a sheaf on Y and it is a sheaf of modules over the sheaf $f_*\mathcal{O}_X$. It is not yet an $\mathcal{O}_Y$-module, but it is very easy to turn it into one: just send a function $\beta \in \mathcal{O}_Y$ along the map $\mathcal{O}_Y \to f_*\mathcal{O}_X$ and then let it act on $f_*\mathcal{F}$. By abuse of notation we refer to this $\mathcal{O}_Y$-module as $f_*\mathcal{F}$. Because f_* concides with the plain sheaf pushforward at the level of sheaves, this abuse of notation is not a big problem. Furthermore f_* is a left exact functor, since that is true already at the level of plain sheaves.

Recall that the sheaf pullback f^{-1} is an exact functor. The pullback of an $\mathcal{O}_Y$-module to Y, however, will not be the plain sheaf pullback "dressed" with extra structure, as the pushforward was. It really is different as a sheaf. Indeed the sheaf pullback $f^{-1}\mathcal{G}$ of an $\mathcal{O}_Y$-module $\mathcal{G}$ is a sheaf on X, but not itself an $\mathcal{O}_X$-module. It is an $f^{-1}\mathcal{O}_Y$-module, and we have a canonical map

$$f^{-1}\mathcal{O}_Y \to \mathcal{O}_X.$$

What we do is define

$$f^*\mathcal{G} = \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}\mathcal{G},$$

which is an $\mathcal{O}_X$-module.

Since $f^*\mathcal{G} \neq f^{-1}\mathcal{G}$ as sheaves, the exactness of f^{-1} does not imply exactness of f^*. And in fact f^* is only right exact in general, essentially because the tensor product is only right exact in general. Also the functors (f^*, f_*) between categories of sheaves of $\mathcal{O}_X$-modules and $\mathcal{O}_Y$-modules, are adjoint functors.

Here is some terminology, which we will probably never use.

Definition 17.2. The morphism f is said to be flat if f^* is exact.

Let R be a commutative ring. We have introduced the spectrum of R in Section **??** in a somewhat informal manner in terms of algebraic sets. A more general and abstract definition is as the set

$$\mathrm{Spec}(R) = \{P \subset R \mid P \text{ is a prime ideal}\}$$

equipped with a topology which we now describe. (Recall that by definition of a prime ideal we allow $P = \{0\}$, but we exclude $P = R$.) For any subset $T \subset R$ we may consider

$$V(T) = \{P \in \mathrm{Spec}(R) \mid P \supset T\}$$

The Zariski topology on $\operatorname{Spec}(R)$ is obtained by declaring all the sets $V(T)$ to be closed [**?**, 10.17]. It turns out that the topology is generated by certain open sets of the form

$$D(f) = \operatorname{Spec}(R)\backslash V(\{f\}).$$

Remark 17.1. *In fact* $\operatorname{Spec}(R_f) \cong D(f)$.

Let $X = \operatorname{Spec}(R)$. We wish to think of R as the ring of functions on X. More precisely we wish to define a sheaf $\mathcal{O}_X$ with $\Gamma(X, \mathcal{O}_X) = R$. We define $\mathcal{O}_X$ by setting

$$\mathcal{O}_X(D(f)) = R_f = \{a/f^N \mid a \in R \text{ and } N \geq 0\}/\sim$$

the localisation (for $f \neq 0$). Since these sets form a base, this assignment determines the sheaf completely. It is just required to check that this assignment actually defines a sheaf.

If $D(g) \subset D(f)$ then we must assign a morphism $R_f \to R_g$. This involves a little bit of commutative algebra. If an element $a \in R$ is outside the union of all maximal ideals of R, then a is invertible. Every maximal ideal is prime, so in particular if a is outside the union of all prime ideals then a is invertible. The inclusion $D(g) \subset D(f)$ tells us that the image of f under $R \to R_g$ is not contained in any prime ideal of R_g, hence is invertible. It follows from the universal property of localisation that there exists a morphism $R_f \to R_g$.

If the inverse of f in R_g is a/g^d, then $af - g^d$ is annihilated by some power of g and so $g^e = af$ for some e. The morphism $R_f \to R_g$ is thus explicitly $b/f^n \mapsto ba^n/g^{en}$.

All this shows that we have a presheaf on the base of the Zariski topology. We still need to check that it is a sheaf on the base (and hence will induce a sheaf in the usual sense). Roughly speaking the sheaf-on-a-base condition to be checked is that, for any cover an intersection of two distinguished opens by other distinguished opens, the gluing condition holds. The present case is a little easier than the general case since $D(f) \cap D(g) = D(fg)$, so it suffices to check covers of distinguished opens by other distinguished opens. Such covers are finite, and if $D(f) \subset \bigcup_i D(g_i)$ then $\{g_i\}$ generates the unit ideal in R_f (this is if and only if apparently).

The gluing condition can be phrased as exactness of the sequence

$$0 \to M_f \to \bigoplus_i M_{fg_i} \to \bigoplus_{i,j} M_{fg_ig_j},$$

which in turn is proved from the fact that $\{g_i\}$ generates the unit ideal in R_f. In fact this is the following lemma applied to R_f in place of R:

Lemma 17.2 ([**?**, Lemma 10.2.4]). *Let M be an R-module, and $\{g_i\}$ generating the unit ideal. Then*

$$0 \to M \to \bigoplus_i M_{g_i} \to \bigoplus_{i,j} M_{g_ig_j}$$

is exact.

Just for definiteness we now state

Definition 17.3. The structure sheaf of the affine scheme $X = \mathrm{Spec}(R)$ is by definition the sheaf constructed in the preceding remarks, and is deinted $\mathcal{O}_X$.

The localisation construction, which is fundamental to the definition of $\mathcal{O}_X$, can be performed for modules too: for an R-module M we have an R_f-module $M_f = R_f \otimes_R M$. We define a sheaf of $\mathcal{O}_X$-modules $\widetilde{M}$ by setting $\widetilde{M}(D(f)) = M_f$. Recalling that points of X are prime ideals in R: the stalk $\widetilde{M}_P$ at a prime ideal P winds up being the localisation of M by the multiplicative set $R \backslash P$.

Since we are discussing the Zariski topology and structure sheaves, we might as well go the whole hog and recall the notion of scheme in general. A topological space which has a cover by subspaces homeomorphic to spaces of the form $\mathrm{Spec}(R)$ (called "open affines"), is called a scheme. More or less.

Definition 17.4. Let $\mathcal{F}$ be a sheaf of $\mathcal{O}$-modules on a scheme X. We say that $\mathcal{F}$ is quasi-coherent if one of the following equivalent conditions holds:

- (See [**?**, Definition 13.2.2]) There exists a cover $\{U_i\}$ by affines, such that $\mathcal{F}|_U \cong \widetilde{M}$ for some $\mathcal{O}(U)$-module M.

- There exists a cover $\{U_i\}$ such that for each $U = U_i$ the restriction $\mathcal{F}|_U$ has a presentation of the form
$$\mathcal{O}_U^{\oplus J} \to \mathcal{O}_U^{\oplus I} \to \mathcal{F}|_U \to 0.$$

Think of quasi-coherence as the condition that lets one reduce questions about a sheaf to a finite collection of questions about modules over rings.

Let's check that the tangent sheaf is quasicoherent. The key computation is the following: Let $D \in \mathrm{Der}(R)$ and $a, f \in R$, then we should have

$$D(a) = D(f(a/f)) = D(f)(a/f) + fD(a/f),$$

so

$$D(a/f) = \frac{1}{f}\left(D(a) - (a/f)D(f)\right).$$

Exercise 17.3. *Deduce that* $\mathrm{Der}(R_f) \cong R_f \otimes_R \mathrm{Der}(R)$.

Hint: This computation is more commonly done for Kähler differentials Ω, namely one is to show that
$$\Omega_{R_f/\mathbb{C}} \cong R_f \otimes_R \Omega_{R/\mathbb{C}},$$
which is a special case of the more general statement

$$S^{-1}\Omega_{R/A} \cong \Omega_{S^{-1}R/A}$$

whenever R is an A-algebra and $S \subset R$ is a multiplicative set. See [**?**, p. 328] (FIND A MORE CLASSIC REFERENCE)

18 Cech cohomology

Discussion of Cech complex: Hartshorne III.4.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be an open cover of a topological space X, and let $\mathcal{F}$ be a sheaf on X. For any finite subset $S = \{i_0, \ldots, i_p\} \subset I$ denote by U_S the open set $\cap_{s \in S} U_s$. Now we define the vector space

$$C^p(\mathcal{U}, \mathcal{F}) = \bigoplus_{\#S = p+1} \mathcal{F}(U_S).$$

This is actually the global sections of a sheaf

$$\mathcal{C}^p(\mathcal{U}, \mathcal{F}) = \bigoplus_{\#S = p+1} j_*^S \mathcal{F}|_{U_S}.$$

We define Cech differentials

$$\mathcal{C}^p(\mathcal{U}, \mathcal{F}) \to \mathcal{C}^{p+1}(\mathcal{U}, \mathcal{F}),$$

but to do so we have to fix an order on I. An element $\alpha \in \mathcal{C}^p(\mathcal{U}, \mathcal{F})$ consists of an element

$$\alpha_{i_0, \ldots, i_p} \in \mathcal{F}(U_{i_0} \cap \cdots \cap U_{i_p})$$

for each tuple $i_0 < \ldots < i_p$, and $d\alpha \in C^{p+1}(\mathcal{U}, \mathcal{F})$ will consist of similar elements

$$(d\alpha)_{i_0, \ldots, i_{p+1}} = \sum_{j=0}^{p+1} (-1)^j \operatorname{Res} \alpha_{i_0, \ldots, \widehat{i_j} \ldots, i_{p+1}}$$

Here Res stands for the restriction from $U_{i_0} \cap \ldots \cap \widehat{U}_{i_j} \cap \ldots \cap U_{i_{p+1}}$ to the smaller open set $U_{i_0} \cap \cdots \cap U_{i_{p+1}}$.

This fits into the general derived functor perspective as follows: We replace our sheaf $\mathcal{F}$ by a resolution

$$\begin{array}{c} \mathcal{F} \\ \downarrow \\ 0 \longrightarrow \mathcal{C}^0 \longrightarrow \mathcal{C}^1 \longrightarrow \cdots \end{array}$$

and then hit the resolution, term-by-term with our functor, in this case $\Gamma(X, -)$, obtaining

$$0 \longrightarrow C^0 \longrightarrow C^1 \longrightarrow \cdots$$

which we then take the cohomology. Cech cohomology is the best thing ever. It's what allows you to turn sheaves on a space into concrete invariants like Betti numbers, etc. Of course, not every open cover gives you the right answer (indeed $\{X\}$ is an open cover of X, and this would usually not be expected to work). In general one imagines that some cover by "sufficiently small" open sets will give the right answer. This is quantified by the definition

of a cover being "acyclic for $\mathcal{F}$". In the case of a smooth manifold, opens homeomorphic to balls in $\mathbb{R}^n$ usually suffice, for quasicoherent sheaves in algebraic geometry, a cover for which all opens and their intersections are open affines in X is sufficient [**?**, Lemma 30.2.6].

Now let's do some examples.

ELLIPTIC CURVES AND ELLIPTIC FUNCTIONS

For example let's compute the Cech cohomology of $\mathcal{O}_E$ using the cover $U_1 = E \backslash p_1$ and $U_2 = E \backslash p_2$. The Cech complex has only two terms

$$C^0 \to C^1$$

where $C^0 = \mathcal{O}_E$ and

$$C^0 = j^1_* \mathcal{O}_E|_{U_1} \oplus j^2_* \mathcal{O}_E|_{U_2},$$

and

$$C^1 = j^{1,2}_* \mathcal{O}_E|_{U_1 \cap U_2}.$$

We hit the complex with the global sections functor: By a bit of elliptic function theory we know that a section of $j^1_* \mathcal{O}_E|_{U_1}$ is a linear combination of $\wp(z - p_1)$ and all derivatives, similarly $\wp(z - p_2)$ etc., for the other factor. As for global sections of C^1 we have linear combinations of all the functions just mentioned, as well as one more: $\zeta(z - p_1) - \zeta(z - p_2)$. Now in

$$\Gamma(E, C^0) \to \Gamma(E, C^1)$$

the kernel of the differential consists of multiples of $(1,1)$ (the constant function in both entries). The cokernel consists of multiples of $\zeta(z - p_1) - \zeta(z - p_2)$.

We have just computed that

$$R^n \Gamma(E, \mathcal{O}_E) \cong \begin{cases} \mathbb{C} & n = 0, 1 \\ 0 & i \geq 2. \end{cases}$$

_____

PUT THIS SOMEWHERE ELSE

A digression on singular homology and cohomology. Hatcher, Ch. 2 for the standard case, Chriss-Ginzburg p. 94 for the Borel-Moore case.

Let X be a topological space. An n-chain on X is a finite $\mathbb{C}$-linear formal combination of continuous maps

$$\sigma : \Delta^n \to X.$$

Here Δ^n denotes the n-dimensional simplex and by convention is the set

$$\Delta^n = (x_0, \ldots, x_n) \in \mathbb{R}^n_{\geq 0} \mid \sum x_i = 1$$

The vertices are assigned labels $0, \ldots, n$, the vertex number i being the one in which $x_i = 1$ and all the other coordinates are 0. Now there are "face maps"

$$F_i : \Delta^{n-1} \to \Delta^n$$

defined by sending the n vertices of Δ^{n-1} to the $n+1$ vertices of Δ^n, in order, skipping i, and extending linearly to the interior of the simplex. The boundary of the n-chain σ is the following $(n-1)$-chain:

$$\partial \sigma = \sum_{i=0}^{n} (-1)^i \sigma \circ F_i.$$

We denote by $C_n^{sing}(X)$ the vector space of all singular n-chains. With ∂ is becomes a homological complex called the singular homology complex. The dual vector spaces $C_{sing}^n(X) = C_n^{sing}(X)^*$ similarly form the singular cohomology complex.

An infinite n-chain on X is an infinite $\mathbb{C}$-linear combination

$$\alpha = \sum_{i=0}^{\infty} a_i \sigma_i$$

of maps $\sigma_i : \Delta^n \to X$. Such an infinite chain is said to be "locally finite" if for all $K \subset X$ compact, the set

$$\{i \mid a_i = 0 \text{ and } \sigma_i(\Delta^n) \cap K \neq 0\}$$

is finite. We denote by $C_n^{BM}(X)$ the vector space of locally finite infinite n-chains. The boundary morphism is defined as usual, and one notes that it is well-defined.

Reference for all this.....

Theorem 18.1. *Let X be a nice topological space. Then there are isomorphisms*

$$H_{sing}^i(X, \mathbb{C}) \cong \check{H}^i(X, \underline{\mathbb{C}}_X).$$

If X is a smooth oriented manifold of dimension n then

$$H_i^{BM}(X) \cong H^{n-i}(X),$$

even if X is not compact.

Here is HTT C.2.16. It seems that Verdier duality happens in and between the categories $D^b(Sh(X))$, i.e., the derived categories of sheaves of abelian groups, or of vector spaces over a fixed field, etc. Then

$$Rf_* \circ \mathbb{D}_X \cong \mathbb{D}_Y \circ Rf_!,$$

where the duality functor $\mathbb{D}_X$ by definition is

$$\mathbb{D}_X(M) = RHom(M, \omega_X).$$

For smooth X, it turns out, the object ω_X is isomorphic to $\underline{\mathbb{C}}_X[n]$. We consider $f : X \to *$. Then on the right hand side (after taking cohomology) we get

$$RHom(R\Gamma_c(X, M), \mathbb{C}_{pt}).$$

If we take $M = \underline{\mathbb{C}}_X$ then this becomes $H_c^0(X)^*$ (the sheaf $\mathbb{C}_{pt}$ is nothing more than a copy of the vector space $\mathbb{C}$, it's not going to get any more acyclic). We pick up the other cohomology groups by taking $M = \underline{\mathbb{C}}_X[i]$ (concentrated in degree $-i$, that is), then the RHS is $H_c^i(X)^*$.

On the LHS we get
$$Rf_*(\mathbb{D}(M))$$
If we take $M = \underline{\mathbb{C}}_X[i]$ then one wants to argue that
$$RHom(\underline{\mathbb{C}}_X[i], \underline{\mathbb{C}}_X[n])$$
is given by resolving the $\underline{\mathbb{C}}_X[n]$ (which as a complex sits in degree $-n$, but whose resolution will be spread out over the degrees $-n, \ldots, -1, 0$), and then the Hom from $\underline{\mathbb{C}}_X[i]$ just picks out the term in degree $-i$. At the level of cohomology we then get
$$H^{n-i}(X).$$
So we recover Poincaré duality
$$H^{n-i}(X) \cong H_c^i(X)^*$$
as required.

19 Finiteness of global sections: algebraic proof

The algebraic proof of the Kodaira-Serre theorem requires an understanding of coherent sheaves on $\mathbb{P}^r$, in particular the twist sheaves $\mathcal{O}(n)$. We briefly recall some background.

The line bundle $\mathcal{O}(1)$ on $\mathbb{P}^r$ is given by a trivialisation with section v_i over D_i and transition functions $v_i = (x_i/x_j)v_j$ on overlaps. In particular if we define a section
$$s_i = x_i^{-1}v_i \qquad \text{on the open set } D_i$$
for each $i = 0, \ldots r$, then on overlaps we find
$$s_i = x_i^{-1}(x_i/x_j)v_j = x_j^{-1}v_j = s_j.$$
In other words the s_i are restrictions of a single section s of $\mathcal{O}(1)$. More generally for any $n \in \mathbb{Z}$ we can consider $\mathcal{O}(n)$, which has local sections v_i satisfying $v_i = (x_i/x_j)^n v_j$. If $n \geq 1$ then $\mathcal{O}(n) \cong \mathcal{O}(1)^{\otimes n}$. In the case of the sphere $X = \mathbb{P}^1$, one can check that the tangent bundle is $T_X \cong \mathcal{O}(2)$. The bundle of 1-forms is the dual bundle (has adjoint transition functions) and is isomorphic to $\mathcal{O}(-2)$.

Remark 19.1. *The tangent and cotangent bundles of $\mathbb{P}^r$ in general are of rank r, so are not isomorphic to some $\mathcal{O}(n)$. But one could compute, for example $\omega_{\mathbb{P}^r} = \wedge^r \Omega_{\mathbb{P}^r}$, and ask for what m is it isomorphic to $\mathcal{O}(m)$. Good exercise. If I did it right then $\omega_{\mathbb{P}^2} \cong \mathcal{O}(-3)$ for example.*

Exercise 19.2. *Compute the Cech cohomology of $\mathcal{O}_{\mathbb{P}^1}$ and $\Omega_{\mathbb{P}^1}$.*

In fact it is useful to give another point of view on $\mathbb{P}^r$ and $\mathcal{O}(n)$. Let us start by considering the polynomial ring, with grading

$$S = \mathbb{C}[X_0, \ldots, X_r] = \bigoplus_{k=0}^{\infty} S_k, \qquad \deg(X_i) = 1 \text{ for each } i.$$

The closed sets in $\mathbb{P}^r$ are by definition the vanishing sets $V(I)$ of *homogeneous* ideals $I \subset S$. These vanishing sets $V(I)$, which technically are subsets of $\mathbb{C}^{r+1}$, are all "conical" in the sense that $x \in V(I)$ implies $\alpha x \in V(I)$ for all $\alpha \in \mathbb{C}^\times$. Thus one pictures such sets as "projected" onto charts, or rather one works with equivalence classes.

The basic open sets of the topology are $D(f) \subset \mathbb{P}^r$, corresponding to the complement of $V(f)$ for homogeneous $f \in S$. Such a basic open set is affine, and one can ask what it is the spectrum of. It turns out that

$$D(f) \cong \mathrm{Spec}(S_f)_0,$$

i.e., just as S is $\mathbb{Z}_+$-graded we see that S_f is $\mathbb{Z}$-graded, and so $(S_f)_0$ is some nontrivial ring.

For example $\mathbb{P}^1 = \{[a_0 : a_1]\}/\sim$ and

$$D_0 = \{[a_0 : a_1] \mid a_0 \neq 0\}/\sim = \{[1 : u]\} \cong \mathrm{Spec}\,\mathbb{C}[u],$$

while $\mathbb{C}[X_0, X_1]_{X_0} = \mathbb{C}[X_0^{\pm 1}, X_1]$ and

$$\mathbb{C}[X_0^{\pm 1}, X_1]_0 = \mathbb{C}[X_1/X_0] \cong \mathbb{C}[u].$$

Similar remarks apply to D_1. Now $D_0 \cap D_1 = D(X_0 X_1)$, which is therefore the spectrum of the ring

$$(\mathbb{C}[X_0, X_1]_{X_0 X_1})_0 = \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}]_0 = \mathbb{C}[(X_1/X_0)^{\pm 1}].$$

The morphisms

$$
\begin{array}{ccc}
D_0 & & D_1 \\
& \nwarrow \qquad \nearrow & \\
& D_0 \cap D_1 &
\end{array}
$$

correspond, at the level of rings, to

$$
\begin{array}{ccc}
\mathbb{C}[X_1/X_0] & & \mathbb{C}[X_0/X_1] \\
& \searrow \qquad \swarrow & \\
& \mathbb{C}[(X_1/X_0)^{\pm 1}] &
\end{array}
$$

Now we present a construction of sheaves on $\mathbb{P}^r$. Suppose M is a graded S-module. Then we assign to the open affine $D(f) \subset \mathbb{P}^r$ the $(S_f)_0$-module $(M_f)_0$. This defines a quasicoherent sheaf on $\mathbb{P}^r$, it turns out, which we denote $\widetilde{M}$.

The Serre twist sheaves arise from a brilliantly simple instance of the general construction. Denote $M = S(1)$ the graded S-module in which $M_n = S_{n+1}$ (so it is S "shifted down" a degree). It turns out that the associated quasicoherent sheaf $\widetilde{M}$ is the sheaf of sections of the line bundle which I have been denoting $\mathcal{O}(1)$.

Let's work out the details in an example. On $\mathbb{P}^1$ we have

$$\widetilde{M}(D_0) = (\mathbb{C}[X_0, X_1]_{X_0})_1 = \mathbb{C}[X_0^{\pm 1}, X_1]_1 = \mathbb{C}[X_1/X_0]X_0,$$
$$\widetilde{M}(D_0) = \mathbb{C}[X_1/X_0]X_1,$$

and

$$\widetilde{M}(D_0 \cap D_1) = \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}]_1 = \mathbb{C}[(X_1/X_0)^{\pm 1}]X_0.$$

We may now think of $v_0 = X_0$ as the trivialising section over D_0, with $\Gamma(D_0, \widetilde{M}) = \mathbb{C}[u]v_0$, and similarly $\Gamma(D_1, \widetilde{M}) = \mathbb{C}[x]v_1$ (where $v_1 = X_1$ and $x = X_0/X_1 = 1/u$). Over the intersection $D_0 \cap D_1$ we have $v_1 = (X_1/X_0)v_0$.

Now we return to general notions for a while, before tackling the proof of the Kodaira-Serre theorem. We introduce some important adjectives for modules over a ring.

Definition 19.1. Let R be a commutative ring and M an R-module.

- We say that M is finitely generated if there exists a surjective morphism $R^n \to M$ for some finite n.

- We say that M is a finitely presented R-module if there exists a surjective morphism $\varphi : R^n \to M$ for some finite n, and if there exists a surjective morphism $R^m \to \ker(\varphi)$ for some finite m,

- We say that M is a coherent R-module if there exists a surjective morphism $R^n \to M$ for some finite n, and if for any morphism (surjective or not) $\varphi : R^N \to M$ there exists a surjective morphism $R^m \to \ker(\varphi)$ for some finite m.

If R is a Noetherian ring, then all three of these notions become identical.

Exercise 19.3. *Prove that if R is Noetherian, then every finitely generated R-module is coherent.*

Now let us sketch the relationship between vector bundles and $\mathcal{O}$-modules. A vector bundle over X is a topological space E with a surjection $\pi : E \to X$, a vector space structure on $\pi^{-1}(x) \subset E$ for each $x \in X$, and trivialisations $\pi^{-1}(U) \to U \times \mathbb{C}^n$ for each U in some atlas of X, such that the induced transition maps $U_i \cap U_j \to GL_n(\mathbb{C})$ lie in whatever category of maps is appropriate.

Given a vector bundle E, for every $U \subset X$ we should have a set of "sections" $\Gamma(U, E)$ forming a vector space. Furthermore $\Gamma(U, E)$ should naturally be a module over $\mathcal{O}_X(U)$. So

E gives rise to a presheaf $\mathcal{F}$ of $\mathcal{O}$-modules. It is actually a sheaf, because of the way we have constructed it. If U is one of the charts on which E is trivialised, then the trivialisation induces an isomorphism $\mathcal{F}|_U \cong \mathcal{O}|_U^{\oplus n}$ of sheaves on U. We will call any sheaf of $\mathcal{O}$-modules which is locally of this form: a "locally free" sheaf.

We would like to be able to use tool from homological algebra. But the collection of locally free sheaves does not form an abelian category. Indeed consider a fixed trivial vector bundle $E = X \times \mathbb{C}^n$ and consider a function f on X which vanishes at just one point $p \in X$. Then multiplication by f is a morphism $E \to E$, whose cokernel is a sheaf with support $\{p\}$ – clearly not a locally free sheaf. So it will be necessary to find an abelian category containing the locally free sheaves.

First an excentric definition.

Definition 19.2. Let $\mathcal{F}$ be a quasicoherent sheaf of $\mathcal{O}$-modules on X. We say $\mathcal{F}$ is coherent if there exists a cover of X by affine opens $U_i = \mathrm{Spec}(R_i)$ such that $\mathcal{F}|_{U_i} \cong \widetilde{M_i}$ for a coherent R_i-module M_i.

Now the usual definition.

Definition 19.3. Let $\mathcal{F}$ be a sheaf of $\mathcal{O}$-modules on X. Then $\mathcal{F}$ is said to be coherent if (1) $\mathcal{F}$ is of finite type (i.e., for every point $x \in X$ there is a neighbourhood U for which $\mathcal{F}(U)$ is a quotient of $\mathcal{O}(U)^n$ for some finite n), and (2) for every open $U \subset X$ and finite collection of sections $s_1, \ldots s_N \in \mathcal{F}(U)$, the kernel of the induced map $\mathcal{O}(U)^N \to \mathcal{F}(U)$ is of finite type.

A scheme X is said to be Noetherian if it is "locally Noetherian" and "quasicompact". Intuitively a scheme is Noetherian if it has a finite cover by affine opens $U_i = \mathrm{Spec}(R_i)$ and each R_i is a Noetherian ring. More formally "locally Noetherian" means every point is contained in an open neighbourhood of the form $U \cong \mathrm{Spec}(R)$ where R is Noetherian, and "quasicompact" means every open cover of X has a finite subcover. Yes, this is the definition of compact. Algebraic geometers use the term "quasicompact".

It follows that over a Noetherian scheme, locally free sheaves of finite rank (i.e., sheaves of sections of vector bundles) are coherent sheaves. One can think of the category of coherent sheaves as the "smallest abelian category" containing all the vector bundles.

The following lemma appears as [**?**, II, Exercise 5.5]

Lemma 19.4. *If $f : X \to Y$ is a closed immersion of algebraic varieties, and $\mathcal{F}$ is a coherent sheaf on X, then $f_*(\mathcal{F})$ is a coherent sheaf on Y.*

Proof. In fact the result holds for any finite morphism $f : X \to Y$, so we have to do three things: (1) explain what a finite morphism is, (2) prove that closed immersions are finite morphisms, and (3) prove the lemma.

A morphism of schemes $f : X \to Y$ is said to be finite if there is a cover of Y by affine opens V_i such that each preimage $U_i = f^{-1}(V_i)$ is an open affine, and (writing $V = \mathrm{Spec}(B)$ and $U = \mathrm{Spec}(A)$) the induced map $B \to A$ makes A into a finitely-generated B-module.

To say that $i : X \to Y$ is an inclusion is to say that Y has a cover by open affines $V_i = \mathrm{Spec}(B_i)$, and the intersection $X \cap V_i$ is the vanishing locus of some ideal $I_i \subset B_i$. The corresponding morphism of rings of functions is $B_i \mapsto B_i/I_i$ and the latter is finitely generated (with one generator) over the former.

So inclusions are finite maps.

Let V_i be a cover of Y by open affines, $U_i = f^{-1}(V_i)$, with B_i and A_i as above. We recall that $f_*\mathcal{F}(V) = \mathcal{F}(f^{-1}(V))$ (with the B-module structure on $f_*\mathcal{F}(V)$ given by mapping $B \to A$ and acting on $\mathcal{F}(f^{-1}(V))$). Coherence of $f_*\mathcal{F}$ now reduces to a simple algebraic fact:

If M is a coherent A-module and $B \to A$ a ring homomorphism with A a finitely generated B-module, then M is a coherent B-module.

If this is not actually true... then we just impose a hypothesis that the rings be Noetherian, and then things should work no trouble. $\qquad\qquad\square$

Proof can be found in Hartshorne, Chapter III, section 5.

Theorem 19.5. *Let X be a projective variety and $\mathcal{F}$ a coherent sheaf on X. Then*

$$\dim H^i(X, \mathcal{F}) < \infty \qquad \text{for all } i \geq 0.$$

Proof. By definition our variety X comes with $i : X \to \mathbb{P}^r$ a closed imersion into a projective space. As we have seen in Lemma **??**, the pushforward $i_*\mathcal{F}$ is a coherent sheaf on $\mathbb{P}^r$. Furthermore

$$H^i(X, \mathcal{F}) \cong H^i(\mathbb{P}^r, i_*\mathcal{F}).$$

(this is [**?**, III, Lemma 2.10], but the basic idea is that $H^i(X, -)$ are the derived functors of pushforward to a point, so the key is to verify that pushforwards compose nicely, i.e., $(f \circ g)_* = f_* \circ g_*$) Thus it suffices to prove the statement for $X = \mathbb{P}^r$.

Proposition **??** below says that the result holds for $\mathcal{F}$ of the form $\mathcal{O}(n)$, hence also for finite direct sums of such sheaves.

So let $\mathcal{F}$ be any coherent sheaf on $\mathbb{P}^r$. Firstly we claim that $\mathcal{F}$ is a quotient of a direct sum of sheaves of the form $\mathcal{O}(n_i)$. It follows from a result which is important in its own right: there exists N such that $\mathcal{F} \otimes \mathcal{O}(N)$ is globally generated. A sheaf $\mathcal{G}$ is said to be globally generated if there exists a surjection $\bigoplus_{i \in I} \mathcal{O} \to \mathcal{F}$.

So why does this N exist? Well, look at the cover of $\mathbb{P}^r$ by the standard affine opens D_i (the set where $x_i \neq 0$). Each restriction $\mathcal{F}(D_i)$ is the localisation of a $\mathbb{C}[x_0, \ldots, x_n]$-module M_i which, by hypothesis, is finitely generated. Let $f_i^{(j)}$ be a finite set of generators of M_i. There exists N such that the section $x_i^N f_i^{(j)}$ of $\mathcal{F} \otimes \mathcal{O}(N)$ on D_i is the restriction of a section on $\mathbb{P}^r$. This statement is similar in spirit (and proof) to Lemma 5.3(b) which says that if $\mathcal{F}$ is a quasicoherent sheaf on an affine scheme X and $s \in \mathcal{F}(D(f))$ then there exists N such that $f^N s \in \mathcal{F}(D(f))$ lies in the image of the restriction $\mathcal{F}(X) \to \mathcal{F}(D(f))$. So we have $\mathcal{O}^k \to \mathcal{F} \otimes \mathcal{O}(N)$ for some k, and hence $\mathcal{O}(-N)^k \to \mathcal{F}$.

Since $\mathbb{P}^r$ has a cover by $r+1$ open affine charts, we have $H^i(\mathbb{P}^r, \mathcal{F}) = 0$ for all $i > r$. (Since the cohomology is computed by a Cech cover.)

Thus we write

$$0 \to \mathcal{R} \to \mathcal{E} \to \mathcal{F} \to 0,$$

where $\mathcal{E}$ is a finite direct sum of sheaves $\mathcal{O}(n_i)$. The kernel $\mathcal{R}$ is also coherent (this is a basic and important feature of the notion of coherence). We have a long exact sequence in cohomology

$$\cdots \to H^i(\mathbb{P}^r, \mathcal{E}) \to H^i(\mathbb{P}^r, \mathcal{F}) \to H^{i+1}(\mathbb{P}^r, \mathcal{R}) \to \cdots.$$

For $i = r$ we have $H^{i+1}(\mathbb{P}^r, \mathcal{R}) = 0$, so

$$\cdots \to H^r(\mathbb{P}^r, \mathcal{E}) \to H^r(\mathbb{P}^r, \mathcal{F}) \to 0,$$

and hence finite dimensionality of $H^r(\mathbb{P}^r, \mathcal{F})$. For $i = r-1$ we obtain

$$\cdots \to H^{r-1}(\mathbb{P}^r, \mathcal{E}) \to H^{r-1}(\mathbb{P}^r, \mathcal{F}) \to H^r(\mathbb{P}^r, \mathcal{R}) \to \cdots,$$

and hence we deduce finite dimensionality of $H^{r-1}(\mathbb{P}^r, \mathcal{F})$ from finite dimensionality of $H^r(\mathbb{P}^r, \mathcal{R})$. We continue in this way by induction. $\qquad\square$

See [**?**, III, Theorem 5.1]

Proposition 19.6. *The sheaf cohomology $H = H^i(\mathbb{P}^r, \mathcal{O}(n))$ is as follows: (1) If $0 < i < r$ then $H = 0$, (2) If $n < 0$ and $i = 0$ then $H = 0$, (3)*

20 Third interlude on sheaves: $\mathcal{D}$-modules

For starters this section follows the introduction to [**?**].

For smooth X one can define $\mathcal{D}_X$ inside the sheaf $\mathrm{End}_{\mathbb{C}}(\mathcal{O}_X)$ inductively by $\mathcal{D}_X^0 = \mathcal{O}_X$ and

$$\mathcal{D}_X^{\leq n} = \{\xi \in \mathrm{End}_{\mathbb{C}}(\mathcal{O}_X) \mid [\xi, \mathcal{O}_X] \subseteq \mathcal{D}_X^{\leq n-1}\}.$$

In particular $\mathcal{D}_X^{\leq 0} = \mathcal{O}_X$, and also $T_X \subset \mathcal{D}_X^{\leq 1}$. Indeed if $v \in T_X$ then

$$[v, f](g) = v(fg) - fv(g) = v(f)g.$$

Previously we have discussed the tangent sheaf T_X which is defined by

$$T_X(U) = \mathrm{Der}(\mathcal{O}_X(U)) = \mathrm{Hom}_{\mathcal{O}_X(U)}(\Omega_X(U), \mathcal{O}_X(U)).$$

We checked that this definition works in Exercise **??**, with the outcome that T_X is a quasi-coherent sheaf on X. A similar idea shows that $\mathcal{D}_X$ is a quasicoherent sheaf too.

The general structure of $\mathcal{D}_X$, for smooth X at least, is quite easy to picture. Start with $X = \mathbb{A}^n$, then $\mathcal{D}_X(X)$ consists of linear combinations of expressions of the form

$$f(x_1,\dots,x_n)\left(\frac{\partial}{\partial x_1}\right)^{i_1}\cdots\left(\frac{\partial}{\partial x_n}\right)^{i_n}. \tag{20.1}$$

If we choose different coordinates on X then a given element of $\mathcal{D}_X(X)$ will have an expression of the same general shape in the new coodinates, but the coefficient functions f will be different in general, and change according to the usual rules of calculus. This applies to the restriction maps on a single variety too, so one should just think of $\mathcal{D}_X(U)$ as consisting of expressions as above on each sufficiently small open U.

For example on $\mathbb{P}^1$ with coordinate z and $w = 1/z$, the section $z\partial_z^2$ corresponds, in the opposite chart, to

$$2w^2\partial_w + w^3\partial_w^2.$$

One starts with

$$\partial_z = \frac{\partial w}{\partial z}\partial_w = -\frac{1}{z^2}\partial_w = -w^2\partial_w,$$

and repeats.

Remark 20.1. *In [?] the authors define $\mathcal{D}_X$ to be the sheaf of algebras generated by $\mathcal{O}_X$ and T_X. But in fact $\mathcal{D}_X$ as defined above is generated by these two subsheaves. There is a proof that goes via associated graded, as presented in [?] and [?]. Step 1: Verify that*

$$[\mathcal{D}_X^{\leq p}, \mathcal{D}_X^{\leq q}] \subset \mathcal{D}_X^{\leq p+q-1}.$$

Step 2: As a formal consequence we have a morphism $\psi : \mathrm{Sym}_{\mathcal{O}_X} T_X \to \mathrm{gr}\,\mathcal{D}_X$. We should like to show this morphism is surjective, in fact we will show it is an isomorphism. Step 3: Construct a morphism

$$\phi^n : \mathrm{gr}\,\mathcal{D}_X \to \mathrm{Sym}^n_{\mathcal{O}_X} T_X,$$

assuming it to have been done already for $j < i$, as follows (base case $n = 1$ is OK). Composition

$$\mathcal{O}_X \otimes \mathcal{D}_X^{\leq n} \xrightarrow{[\cdot,\cdot]} \mathcal{D}_X^{n-1} \to \mathrm{gr}^{n-1}\mathcal{D}_X$$

descends to

$$\mathcal{O}_X \otimes \mathrm{gr}^n \mathcal{D}_X \to \mathrm{gr}^{n-1}\mathcal{D}_X \cong \mathrm{Sym}^{n-1}_{\mathcal{O}_X} T_X$$

and is, for each fixed element of gr^n, a derivation on the first factor. So we define ϕ^n as the composition

$$\mathrm{gr}^n \mathcal{D}_X \to \mathrm{Der}(\mathcal{O}_X, \mathrm{Sym}^{n-1}_{\mathcal{O}_X}) \cong T_X \otimes \mathrm{Sym}^{n-1}_{\mathcal{O}_X} T_X \to \mathrm{Sym}^n_{\mathcal{O}_X} T_X.$$

Step 4: If $\phi^n(P) = \sum \xi \otimes P' \in T_X \otimes \mathrm{Sym}^{n-1}_{\mathcal{O}_X}$ then one confirms that $\sum P' \circ \xi \equiv P$ in the associated graded.

A $\mathcal{D}$-module on X means a sheaf of modules over the sheaf of algebras $\mathcal{D}_X$-modules. Since the algebras $\mathcal{D}_X(U)$ are not commutative in general, one has to distinguish left and right $\mathcal{D}$-modules.

The canonical example of a left $\mathcal{D}_X$-module is $\mathcal{O}_X$ with the action which defines $\mathcal{D}_X$. An interesting and important example of a right $\mathcal{D}_X$-module (on a smooth n-dimensional variety) is the sheaf of n-forms ω_X. Consider an open chart U with coordinates $x_1, \ldots, x_n$, so that sections of ω_X over U are expressible as

$$f\eta = f(x_1, \ldots, x_n)\, dx_1 \wedge \ldots \wedge dx_n.$$

The right action of $\mathcal{D}_X(U)$ is characterised by

$$\eta \partial_i = 0, \qquad i = 1, \ldots, n.$$

The requirement of being a right $\mathcal{D}_X$-module now determines the action on all $f\eta$, indeed

$$(f\eta)\partial_i = \eta(f\partial_i) = \eta([f, \partial_i] + \partial_i f) = -f_i\eta.$$

It is a good exercise to check that this definition makes sense, and to feel why the construction yields a right $\mathcal{D}_X$-module and not a left $\mathcal{D}_X$-module. The construction can be expressed more invariantly as

$$(s)X = -\mathcal{L}_X(s)$$

where X is any vector field and $\mathcal{L}_X$ is the Lie derivative, here s is any section of ω_X.

In fact there is a close relation between left and right $\mathcal{D}_X$-modules. Let M be a left $\mathcal{D}_X$-module, then we can produce a corresponding right $\mathcal{D}_X$-module

$$M \otimes_{\mathcal{O}} \omega$$

with action

$$(m \otimes \eta)\partial = -(\partial m) \otimes \eta - m \otimes (\partial \eta).$$

Let's work more concretely for a moment. Let $X = \mathbb{A}^1$, and write the algebra $\Gamma(X, \mathcal{D}_X)$ as $D = \mathbb{C}[x, \partial]$. This is not really a polynomial algebra, it is the free associative on generators x and ∂ module the ideal of relations generated by

$$\partial x - x\partial = 1.$$

By a variant of the PBW theorem D has a basis consisting of elements the form $x^m \partial^p$ for $m, p \in \mathbb{Z}_+$. Now we can consider all sorts of submodules and quotient modules, for example $D\partial \subset D$ is a left submodule (isomorphic to D itself), and we can consider the quotient

$$D/D\partial$$

which has a basis $\{x^m \mid m \in \mathbb{Z}_+\}$. We see that $D/D\partial = \mathbb{C}[x] = \mathcal{O}_X(X)$. On the other hand we can consider

$$D/Dx = \langle \partial^p \mid p \in \mathbb{Z}_+ \rangle.$$

We denote the image of $1 \in D$ in this module by the symbol δ. This is appropriate because it satisfies the relation $x\delta = 0$, and so can be conceptualised as the Dirac delta function.

DRAW A PICTURE HERE.

Let's stick with $X = \mathbb{A}^n$ for a while, and write $D = \mathbb{C}[x_1, \ldots, x_n, \partial_1, \ldots, \partial_n]$ as before. We may describe systems of linear PDEs in the language of $\mathcal{D}$-modules, and even think about what it would mean to solve them "algebraically". Let $u_1, \ldots, u_\ell$ be ℓ unknown functions of the variables $x_1, \ldots, x_n$ which satisfy a system of k partial differential equations

$$\sum_{j=1}^{\ell} P_{ij} u_j = 0, \qquad i = 1, \ldots, k \tag{20.2}$$

Let's consider the u_j as *formal* objects at first, and consider the free D_X-module D_X^ℓ with generators $\widetilde{u}_1, \ldots, \widetilde{u}_\ell$. So vectors here are just linear combinations of objects for the form (**??**) applied to the u_i. An assignment of *actual* functions $u_1, \ldots, u_\ell$ extends to a morphism

$$D_X^\ell \to \mathcal{O}_X, \qquad \widetilde{u}_i \mapsto u_i.$$

Saying that the functions satisfy the system (**??**) says in particular that each vector

$$\sum_{j=1}^{\ell} P_{ij} \widetilde{u}_j \in D_X^\ell$$

is sent to zero in $\mathcal{O}_X$. So in fact we have a morphism

$$\varphi : D_X^k \to D_X^\ell$$

defined by $\varphi(\widetilde{v}_i) = \sum_{j=1}^{\ell} P_{ij} \widetilde{u}_j$, and the quotient D_X-module $M = D_X^\ell / \varphi(D_X^k)$. Then a solution to the system of PDEs is the same thing as a morphism $M \to \mathcal{O}_X$ of D_X-modules. In algebraic language, we are to study

$$\mathrm{Hom}_D(M, \mathcal{O})$$

for finitely presented D-modules M.

Now, if we want to work really algebraically, then $\mathcal{O}$ means the ring $\mathbb{C}[x_1, \ldots, x_n]$ of polynomial functions (which we can relax to rational functions if we restrict to Zariski open sets $U \subset \mathbb{A}^n$). But in this ring even the simplest of differential equations often have no solution. For example $n = 1$ and

$$P = P_{11} = \partial_x - 1$$

The solutions (in C^∞ let's say) are Ae^x, these are all the solutions in C^∞, and there are no solutions in $\mathbb{C}[x]$.

For another example we take $X = \mathbb{A}^1 \backslash \{0\}$ (i.e., the spectrum of $\mathbb{C}[x, x^{-1}]$) and $P = \partial_x - (2/3)x^{-1}$, then in simply connected domains away from $x = 0$ we have the C^∞ solutions $Ax^{2/3}$. The thing about this solution is that if we analytically continue around the origin,

say from a small disc centred on $x = 1$, anticlockwise around the unit circle (we're in $\mathbb{A}^1$, the complex plane) back to the starting point, the solution comes back multiplied by $e^{2\pi i/3}$. In particular there is no globally well-defined solution on X, *even* working in C^∞. The lesson of this example is that we ought to be open to the idea of sheaves of solutions, not just solutions on the whole space.

Is there a clean way to describe this locally constant sheaf, simpler than "multiples of the analytic function $x^{2/3}$"

The Riemann-Hilbert correspondence is deeply transcendental. From HTT p. 99:

> Although our objectives in this book are algebraic D-modules (D-modules on smooth algebraic varieties), we have to consider the corresponding analytic D-modules (D-modules on the underlying complex manifolds with classical topology) in defining their solution (and de Rham) complexes.

Also the RH correspondence involves local systems, and local systems in the Zariski topology are all trivial (reference for this??). At the same time, the de Rham functor between derived categories

$$DR_X(M^\bullet) = \Omega_X \otimes^L_{D_X} M^\bullet$$

apparently makes sense either analytically or algebraically (though HTT definitely make a point of working analytically here). So a priori it does seem at least plausible that in the derived category analytically locally constant sheaves could be represented by purely Zariski type objects.

Random notes from HTT.

A $\mathbb{C}_X$-module is the same thing as a sheaf of vector spaces. So $D^b(\mathbb{C}_X)$ denotes the derived category of sheaves of vector spaces.

Def 4.2.2. A local system is a sheaf which locally looks like a constant sheaf of fin dim vector spaces. In other words a locally free $\mathbb{C}_X$-module of finite rank.

Observe that a vector bundle is essentially a locally free $\mathcal{O}_X$-module of finite rank. But the notion is extremely different.

A constructible sheaf is a complex of sheaves of vector spaces whose cohomology sheaves are local systems.

To start with DR_X lands in $D^b(\mathbb{C}_X)$. Later it is shown to land in constructible sheaves. This is Kashiwara's constructibility theorem [**?**, Theorem 4.6.3]. Then Thm 4.6.6 is the claim that the restriction to honest D-modules lands in perverse sheaves. Apparently the proof of Kashiwara's theorem is easier for algebraic D-modules, but the definition of the functors is still to first analytify (see p. 121.). Also Remark 4.7.3 emphasises that algebraic $\Omega_X \otimes^L_{D_X} M^\bullet$ is not a good object.

The algebraic case of Kashiwara is Thm 4.7.7. The proof is shorter, but uses etale morphisms and doesn't look necessarily easier....

Much later, after discussion of regular D-modules, we come to Theorem 7.2.2

$$DR_X : D^b_{rh}(D_X) \to D^b_c(X)$$

is an equivalence of categories, and (Theorem 7.2.5) the pure D-modules go to the perverse sheaves.

TO DO: $\partial_x - \lambda$ is not regular. Is $x\partial_x - \lambda$ regular?

A nice narrative here is that we have to have a foot in both complex analytic and (projective) algebraic worlds. We only get enough solutions by working analytically (example $D/D(\partial_x - \lambda)$ and $D/D(x\partial_x - \lambda)$ have no algebraic solutions in general). But on the other hand we only get an equivalence to $D^b_c(X)$ by restricting from the zoo of all analytic differential equations to the regular ones, and regularity cannot be defined analytically.

The note of `sanathd` is quite nice. Describes vector bundles with integrable connection as the simplest case of holonomic. Also the RH correspondence is super simple in principle in this case: at least the inverse functor is, just take the local system and tensor with $\mathcal{O}_X$ to get a locally free $\mathcal{O}_X$-module (i.e., vector bundle!), also d comes from $d : \mathcal{O}_X \to \Omega_X$ tensored with the identity on the local system.

A nice way to think of the RH equivalence is that the the DR functor sends many D-modules to the same local system, but among all of these there is a sort of "distinguished" one, the one with regular singularities. How is this special D-module constructed from the local system (in general perverse sheaf)

21 Integrable connections and local systems

There is a fundamental notion in differential geometry of a rank n vector bundle E with a connection ∇, which is to say a specification of how we should "differentiate" sections of E, or which is the same thing, a specification of which sections of E we are to single out as being "horizontal" in the same way that the function 1 is declared to be horizontal among all C^∞ functions.

Formally, a connection ∇ on a vector bundle E is a morphism of sheaves of sections

$$\nabla : \mathcal{E} \to \Omega^1_X \otimes_{\mathcal{O}_X} \mathcal{E},$$

or equivalently for each $\xi \in T_X(U)$, a morphism

$$\nabla_\xi : \mathcal{E}(U) \to \mathcal{E}(U)$$

satisfying

  (i) $\nabla_{f\xi}(s) = f\nabla_\xi(s)$,

  (ii) $\nabla_\xi(fs) = \xi(f)s + f\nabla_\xi(s)$.

Equivalently
$$\nabla(fs) = df \otimes s + f \cdot \nabla(s).$$
Such a map is *not* $\mathcal{O}_X$-linear. Such a map extends to a sequence of maps
$$d : \Omega_X^k \otimes_{\mathcal{O}_X} \mathcal{E} \to \Omega_X^{k+1} \otimes_{\mathcal{O}_X} \mathcal{E}.$$
A special class of connections are those for which $d \circ d = 0$, these are called *integrable connections*. This condition cane be expressed in terms of ∇_ξ, and says that:

(iii) $\nabla_{[\xi_1, \xi_2]}(s) = [\nabla_{\xi_1}, \nabla_{\xi_2}](s).$

Now integrable connections form a special class of $\mathcal{D}_X$-modules. Concretely, if $(\mathcal{E}, \nabla)$ is an integrable connection, then

$$f \cdot s \mapsto fs, \qquad \xi \cdot s = \nabla_\xi(s)$$

uniquely defines a structure of $\mathcal{D}_X$-module on $\mathcal{E}$. To check that this is true, one must observe (1) that $\mathcal{O}_X$ and T_X generate $\mathcal{D}_X$, and (2) the relations in $\mathcal{D}_X$ are satisfied by this action, this in turn follows from the axioms (i), (ii) and (iii) defining the notion of integrable connection.

Now let's consider a very simple construction that can be performed on $(\mathcal{E}, \nabla)$: we consider the subsheaf
$$\ker(\nabla) \subset \mathcal{E}.$$
Just as ∇ is a morphism of sheaves but not a morphism of $\mathcal{O}_X$-modules, the sheaf $\ker(\nabla)$ is not an $\mathcal{O}_X$-module in any natural way. That's OK, it doesn't have to be. It is simply a sheaf of vector spaces.

But in fact $\ker(\nabla)$ is not just a random sheaf of vector spaces, it is in general (and for this we have to pass to the complex analytic topology) a locally constant sheaf. Indeed let $x \in X$ and $U \subset X$ a neighbourhood of x homeomorphic to a ball in $\mathbb{C}^n$ over which E is trivialised. Any section s is identified with a tuple $(u_1, \ldots, u_n)$ of holomorphic functions over U and the condition $s \in \ker(\nabla)$

Definition 21.1. Let X be a C^∞ manifold or complex manifold. A distribution on X is a sub- vector bundle $D \subset T_X$. The distribution is said to be involutive if $[\xi, \eta] \in D$ for all sections $\xi, \eta \in D$.

If we consider the total space E of our vector bundle (whose sheaf of sections we are denoting $\mathcal{E}$), then the connection ∇ defines a distribution on E

Theorem 21.1 (Frobenius)**.** *If D is an involutive distribution of rank r on a manifold E and $p \in E$, then there exists a neighbourhood U of p in E and a submanifold $S \subset U$ of dimension r containing p such that at each point $s \in S$ we have $D_s = T_s S \subset T_s E$.*

We can deduce from this theorem, the following

Theorem 21.2 (Frobenius). *Let $\pi : E \to X$ be a vector bundle with integrable connection over a smooth manifold X, and let $x \in X$ and $p \in \pi^{-1}(x)$. Then there exists a neighbourhood U of x in X, for which there is a unique section $s \in \mathcal{E}(U)$ satisfying $\nabla(s) = 0$ and $s(x) = p$.*

In particular, choosing a basis $\{p_1, \ldots, p_n\}$ of E_x, and making the identifications

$$\mathbb{C}^n \to E_x \to \mathcal{E}(U),$$

equips $\ker(\nabla)$ with the structure of a local system.

Conversely, given a local system $\mathcal{L}$ it is possible, in fact very easy, the construct a vector bundle with integrable connection. We let

$$\mathcal{E} = \mathcal{L} \otimes_{\underline{\mathbb{C}}_X} \mathcal{O}_X,$$

which is a vector bundle, because on open charts in which $\mathcal{L}(U)$ is identified with some vector space V, clearly $\mathcal{E}(U)$ is similarly identified with $V \otimes \mathcal{O}_X(U)$. The transition maps between different such charts are just constant matrices. Finally we can endow $\mathcal{E}$ with a connection by defining, locally,

$$\nabla(v \otimes f) = v \otimes df.$$

To put this another way, there is a natural inclusion

$$\mathcal{L} \subset \mathcal{E}$$

and the connection is constructed in such a way that $\mathcal{L} \subset \ker(\nabla)$.

Thus it is basically clear already that if we start from a local system $\mathcal{L}$ and perform the operations $(-) \otimes_{\underline{\mathbb{C}}_X} \mathcal{O}_X$ and then $\ker(\nabla)$, we will recover $\mathcal{L}$ up to isomorphism. Similarly it is not hard to show that if we start from a vector bundle with integrable connection (E, ∇) and apply the operations $\ker(\nabla)$ and then $(-) \otimes_{\underline{\mathbb{C}}_X} \mathcal{O}_X$, we will recover (E, ∇) up to isomorphism.

This equivalence between integrable connections and local systems is clean and lovely. But more general versions are needed in applications. As we saw above the differential equation $(x\partial - 2/3)u = 0$, which corresponds to the $\mathcal{D}$-module $M = D/D(x\partial - 2/3)$, has a sheaf of solutions on $\mathbb{C}$ which is not quite a local system, and indeed M itself is not quite a vector bundle with integrable connections, there is something singular about the connection at $x = 0$.

The correspondence described here is generalised by the so-called Riemann-Hilbert correspondence. It treats $\mathcal{D}$-modules satisfying appropriate finiteness conditions and relates them to *constructible sheaves*. It is a statement about derived categories. It takes a lot of work to properly understand the Riemann-Hilbert correspondence and its proof. But perhaps the following observation could help one get started. The correspondence described above was simply $(\mathcal{E}, \nabla) \mapsto \ker(\nabla)$. We have seen that ∇ is naturally the start of a complex of sheaves

$$0 \to \mathcal{E} \to \Omega_X^1 \otimes_{\mathcal{O}_X} \mathcal{E} \to \Omega_X^2 \otimes_{\mathcal{O}_X} \mathcal{E} \to \ldots \to \Omega_X^n \otimes_{\mathcal{O}_X} \mathcal{E} \to 0,$$

and $\ker(\nabla)$ is simply the cohomology in degree 0 of this complex. One can imagine trying to take the other cohomology groups into account also. Natural questions begin to suggest themselves immediately: does one expect the higher cohomology groups to be local systems, or local systems on a large open set in some precise sense? If so, in what sense?) Packaging the resulting data into some sort of derived object of the category of constructible sheaves.

Note to self: The proof of RH comes after a very prolonged throat-clearing about six-functor formalisms and derived categories. While, yes, it's true and important that the RH correspondence respects the six functors, is it perhaps possible to skip all of that and just "do" it, as we did above for the case of local systems?

22   The hypergeometric functions

Taken from the book of Ince [**?**, p. 389].

Riemann considered the following problem: find a differential operator $Lin\Delta_1$, such that

- the ODE $Lf = 0$ is order 2,

- the solutions are holomorphic at all points in the complex plane except three: a, b and c,

- There exist two constants α_1, α_2, and two solutions defined in a neighbourhood of $z = a$ of the form
  $$(z-a)^{\alpha_1} g_1(z) \qquad (z-a)^{\alpha_2} g_2(z)$$
  where g_1 and g_2 are holomorphic in the neighbourhood such that

23   Singular support and holonomic $\mathcal{D}$-modules

Above we commented on the order filtration on $\mathcal{D}$ and its associated graded $\operatorname{gr} \mathcal{D} \cong \operatorname{Sym}_{\mathcal{O}_X} T_X$. The associated graded is a big commutative algebra, which means we can ask what it corresponds to geometrically. The answer is that polynomials in elements of T_X are something like polynomial functions on the total space of T^*X. On the other hand $\operatorname{Sym}_{\mathcal{O}_X} T_X$ is a sheaf, not on the total space of T_X^* but rather on X. In fact if $\pi : T_X^* \to X$ is the natural projection,
$$\operatorname{Sym}_{\mathcal{O}_X} T_X \cong \pi_* \mathcal{O}_{T_X^*}.$$

Let M be a $\mathcal{D}_X$-module with a good filtration. Then we can form
$$\mathrm{Ch}_M = \operatorname{supp}\left(\mathcal{O}_{T_X^*} \otimes_{\mathcal{O}_{\pi^{-1}\pi_* T_X^*}} \pi^{-1} \operatorname{gr} M \right),$$
a closed subset in the total space T_X^*. This is a lot to swallow at once, but if
$$M = D/(DP_1 + \cdots DP_k)$$

then Ch_M is the vanishing set of the ideal

$$\{\sigma(Q) \mid DP_1 + \cdots DP_k\}.$$

Here $\sigma(Q)$ is the "principal symbol" of a differential operator.

24   Sheaves of $\mathcal{D}$-modules

Let $f : X \to Y$ be a morphism. We wish to define pushforward and pullback of $\mathcal{D}$-modules along f. Much as in the case of $\mathcal{O}$-modules, we adopt the view that the correct notions will arise by combining the plain sheaf notions with some understanding of the relations between $\mathcal{D}_X$ and $f^{-1}\mathcal{D}_Y$, and between $f_*\mathcal{D}_X$ and $\mathcal{D}_Y$.

First we define pullback as it is easier. Let $\mathcal{G}$ be a left $\mathcal{D}_Y$-module. Consider the $\mathcal{O}$-module pullback

$$f^*\mathcal{G} = \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}\mathcal{G}.$$

We equip it with a left $\mathcal{D}_X$-module structure via the formula

$$\theta(\alpha \otimes s) = \theta(\alpha) \otimes s + \alpha \sum_i \theta(y_i \circ f) \otimes \frac{\partial}{\partial y_i} s. \tag{24.1}$$

Where did the weird looking definition (**??**) come from? The idea is to imagine a solution of the PDE corresponding to $\mathcal{G}$, write down some differential equations that the pullback of such a function should satisfy, and then code them into a $\mathcal{D}$-module. It's really just the chain rule:

$$\frac{\partial s}{\partial x} = \sum_i \frac{\partial y_i}{\partial x} \frac{\partial s}{\partial y_i}.$$

We can do this construction in a slightly more fancy way, which will lead us to the correct definition of pushforward also. The idea is to consider the $\mathcal{O}_X$-module pullback sheaf on X:

$$f^*\mathcal{D}_Y = \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}\mathcal{D}_Y,$$

which itself carries the nontrivial $\mathcal{D}_X$-module structure (**??**). We denote this sheaf by the symbol $\mathcal{D}_{X \to Y}$, it is a $\mathcal{D}_X$-module, acting on the left. It is also a right $f^{-1}\mathcal{D}_Y$-module.

The D-module pullback can now be defined as

$$f^*\mathcal{G} = \mathcal{D}_{X \to Y} \otimes_{f^{-1}\mathcal{D}_Y} f^{-1}\mathcal{G}.$$

We check that as an $\mathcal{O}$-module we recover

$$f^*\mathcal{G} = \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}D_Y \otimes_{f^{-1}D_Y} f^{-1}\mathcal{G} = \mathcal{O}_X \otimes_{f^{-1}\mathcal{O}_Y} f^{-1}\mathcal{G} = f^*\mathcal{G}.$$

Notice, then, that the pullback of $\mathcal{D}$-modules is right exact because at the level of sheaves it coincides with the pullback of $\mathcal{O}$-modules, which is a right exact functor.

Let us now work out the pushforward of $\mathcal{D}$-modules. Let $\mathcal{F}$ be a *right* $\mathcal{D}_X$-module. Tensoring it with $\mathcal{D}_{X \to Y}$ gives an $f^{-1}\mathcal{D}_Y$-module

$$\mathcal{F} \otimes_{\mathcal{D}_X} \mathcal{D}_{X \to Y}.$$

Notice that here we really are forced to start with a right $\mathcal{D}_X$-module. The result is still a sheaf on X.

This sheaf is still an $\mathcal{O}_X$-module. Is that relevant? Comment on why the $\mathcal{O}_X$-module structure survives.

We can now push it forward to Y, obtaining

$$f_*(\mathcal{F} \otimes_{\mathcal{D}_X} \mathcal{D}_{X \to Y}),$$

which is a right $\mathcal{D}_Y$-module. Why is it a right $\mathcal{D}_Y$-module? Well, it makes sense that is be a $f_* f^{-1}\mathcal{D}_Y$-module, whatever that is. But here the adjuction works in our favour. From

$$\operatorname{Hom}(f^{-1}\mathcal{G}, \mathcal{F}) \cong \operatorname{Hom}(\mathcal{G}, f_*\mathcal{F})$$

we get (putting $f_*\mathcal{G}$ in place of $\mathcal{F}$)

$$\operatorname{Hom}(\mathcal{G}, f_* f^{-1}\mathcal{G}) \cong \operatorname{Hom}(f^{-1}\mathcal{G}, f^{-1}\mathcal{G}),$$

and the unit morphism on the right hand side corresponds to some canonical morphism $\mathcal{G} \to f_* f^{-1}\mathcal{G}$. We use $\mathcal{D}_Y \to f_* f^{-1}\mathcal{D}_Y$ to give $f_*(\mathcal{F} \otimes_{\mathcal{D}_X} \mathcal{D}_{X \to Y})$ a right $\mathcal{D}_Y$-module structure.

This definition involves a composition of a right exact functor $\otimes$ with a left exact functor f_*. The resulting composition is a perfectly well-defined functor, but it is generally not exact on either side. This is a slightly new situation, the majority of functors we study are either left or right exact, and we study their derived functors. This functor does not have derived functors in the usual sense, and to study it using homological algebra we need to use the heavier machinery of derived categories (or, which is basically equivalent, we need to use spectral sequences to compute it in general).

We will use the following notation for the pushforward of $\mathcal{D}$-modules: If $f : X \to Y$ and $\mathcal{F}$ is a $\mathcal{D}_X$-module, then its pushforward is the $\mathcal{D}_Y$-module

$$\int_f \mathcal{F} = f_*(\mathcal{F} \otimes_{\mathcal{D}_X} \mathcal{D}_{X \to Y}).$$

The same notation will also be used (following [?, p. 40]) for the corresponding "derived object"

$$\int_f \mathcal{F} = Rf_*(\mathcal{F} \otimes_{\mathcal{D}_X}^L \mathcal{D}_{X \to Y}).$$

One important case of $\int$, for which we do not need the whole machinery of derived categories to understand, is $\int_i$ for $i : X \to Y$ a closed embedding. It is known that $\int_i$ is exact (see

[**?**, p. 69]). Locally the image $i(X)$ is described by the vanishing of functions in some ideal $I_X \subset \mathcal{O}_Y$, and the local description of $\int_i D_X$ is

$$D_Y/D_Y I_X.$$

This includes the case of the $\delta(x)$-function, as the case $i : * \to Y = \mathbb{A}^1$ embedding a point. Then we can consider $\mathcal{O}_*(*) = \mathbb{C}$ and $\mathcal{D}_*(*) = \mathbb{C}$ as well as $\mathcal{O}_Y(Y) = \mathbb{C}[x]$ and $\mathcal{D}_Y(Y) = D = \mathbb{C}[x, \partial]$. The direct image $\int_i \mathcal{D}_X$ is a skyscraper sheaf at 0, whose global sections is the D-module D/Dx.

Another very important case is that in which X is smooth projective and $p : X \to *$. Then the $\mathcal{D}$-module pushforward along p is essentially de Rham cohomology. More precisely $\int_p \omega_X$ is an object in the derived category of $\mathcal{D}_*$-modules (which is to say $\mathbb{C}$-modules), whose cohomology it turns out is the singular cohomology $H^*(X)$ of X.

Indeed $p^{-1}(\mathcal{D}_*) = \underline{\mathbb{C}}_X$ the constant sheaf on X, and $D_{X\to *} = \mathcal{O}_X \otimes_{\underline{\mathbb{C}}_X} \underline{\mathbb{C}}_X = \mathcal{O}_X$. (I'D LIKE TO FIND AN EXPLICIT REFERENCE FOR THIS CALCULATION.)

As we see in [**?**, ?????] the association

$$\mathcal{F} \mapsto \mathcal{F} \otimes_{D_X}^L \mathcal{O}_X$$

has a name, it is the de Rham functor (closely related to the solution functor).

Anyway, spell out the story of how doing derived $(-) \otimes_D \mathcal{O}$, followed by derived global sections, yields singular cohomology and in the process the Hodge diamond.

Let us do the example of the smooth elliptic curve $E = \mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau$, which is particularly easy to work out concretely since ω_E has a nonvanishing global section dz and hence $\omega_E \cong \mathcal{O}_E$. We have a resolution of $\mathcal{O}_E$ by left $\mathcal{D}_E$-modules which is basically $\mathcal{D} \to \mathcal{D}$ with $f(z)\partial_z^k \mapsto f(z)\partial_z^{k+1}$.

We tensor over $\mathcal{D}$ with the right $\mathcal{D}$-module ω, obtaining

$$\omega \otimes_{\mathcal{D}} \mathcal{D} \to \omega \otimes_{\mathcal{D}} \mathcal{D}.$$

Now this is just

$$\omega \to \omega,$$

with the map being

$$dz f(z) \mapsto -dz f'(z).$$

Now we may take a Cech resolution of both terms, and we obtain a spectral sequence whose first page looks as follows

PICTURE

The horizontal maps are just derivatives. The representative 1 of H^0 gets killed by the derivative, and we get constributions to bidegree $(0,0)$ and $(1,0)$. Similarly the representative ζ of H^1, upon application of derivative becomes a combination of $\wp$, which are zero in Cech cohomology. So ζ contributes to bidegrees $(0,1)$ and $(1,1)$. In total, therefore, we recover $H^*(\int_p \omega_E) = (\mathbb{C}, \mathbb{C}^2, \mathbb{C})$, which are the Betti numbers of E.

elliptic-cech.png

25 Sato description of diagonal pushforward

Do we really need to?

Explain the intuition for $\mathcal{D}$-module pushforward along closed embeddings.

26 Beilinson-Bernstein localisation

Let X be a variety which carries an action of the Lie group G. Then we obtain a linear map $\mathfrak{g} \to T_X(U)$ for each open $U \subset X$. We recall how this map was constructed, intuitively. If we denote the image of $a \in \mathfrak{g}$ by ξ_a then for $x \in X$ and a function f defined near x, we have by definition

$$[\xi_a f](x) = \frac{d}{dt} f(e^{-ta}x)|_{t=0}.$$

In Section **??** above we worked out explicitly the case of $X = G/B$ for $G = SL_2$, with the result

$$E \mapsto -\frac{\partial}{\partial u}$$
$$H \mapsto -2u\frac{\partial}{\partial u} \qquad\qquad (26.1)$$
$$F \mapsto u^2\frac{\partial}{\partial u},$$

for functions of u on the big cell $X_\sigma \subset G/B$, parametrised by

$$u \mapsto \begin{pmatrix} u & -1 \\ 1 & 0 \end{pmatrix}.$$

Back in the general context, by taking arbitrary products we obtain a homomorphism of associative algebras $U(\mathfrak{g}) \to \mathcal{D}_X(U)$ for each U. If we wish to express this in sheaf language, we would say that there is morphism of sheaves of algebras

$$\underline{U(\mathfrak{g})} \to \mathcal{D}_X$$

(where we have the constant sheaf on the LHS).

A quick remark on twisted differential operators. Let E be any $\mathcal{O}$-module on X. Then we can imitate the definition of $\mathcal{D}_X$ to obtain a sheaf $\mathcal{D}_X^E$ defined as the union of

$$\mathcal{D}_X^{E,\leq n} = \{\xi \in \mathrm{End}_{\mathbb{C}}(E) \mid [\xi, \mathcal{O}_X] \subseteq \mathcal{D}_X^{E,\leq n-1}\}.$$

This is called the sheaf of differential operators on E, and in general sheaves of this form are called twisted differential operators (TDO).

We have actually worked out some examples of TDOs already. Recall the line bundles $\mathcal{L}_\Lambda$ on $X = G/B$ which we studied in the Borel-Weil theorem. Recall that from their construction in the form $E = G \times M/\sim$ over $X = G/B$, the total spaces of these bundles acquire a compatible G-action. We say they are G-equivariant vector bundles. We interpreted sections s of E as functions $G \to M$ satisfying a certain equivariance condition. At the level of such functions, the action of G is

$$[gf](x) = f(g^{-1}x).$$

We can translate this action back into an action on sections. If s is a section of E then we discover that gs is given by the formula

$$[gs](xB) = gs(g^{-1}xB).$$

Now we check that the following formula defines an action of $\mathfrak{g}$ by differential operators on sections of E.

$$[\xi_a s](x) = \frac{d}{dt}e^{ta}s(e^{-ta}x)|_{t=0}.$$

109

We did this already, and for the line bundle $\mathcal{L}_{-k\varpi}$ we got the result

$$E \mapsto -\frac{d}{du},$$
$$H \mapsto -2u\frac{d}{du} - k, \tag{26.2}$$
$$F \mapsto u^2\frac{d}{du} + ku.$$

The abstract statement here is that if X is a space with an action of G and E is a G-equivariant vector bundle on X then there is a morphism

$$\underline{U(\mathfrak{g})} \to \mathcal{D}_X^E.$$

Everything we do below can be done for general vector bundles, but we will work primarily with the trivial line bundle.

Let V be a left $U(\mathfrak{g})$-module now. We define a sheaf $\mathcal{V}$ of $\mathcal{D}_X$-modules on X as the quasicoherent sheaf associated with $\Gamma(X, \mathcal{D}_X) \otimes_{U(\mathfrak{g})} V$, that is

$$\mathcal{V}(U) = \mathcal{D}_X(U) \otimes_{U(\mathfrak{g})} V.$$

This is called *Beilinson-Bernstein localisation*. We shall often denote the $\mathcal{D}$-module associated with V as $\Delta(V)$.

Let's consider $\mathfrak{g} = sl_2$ and the case of the trivial $\mathfrak{g}$-module $V = \mathbb{C}$. We observe that $V = U(\mathfrak{g})/U(\mathfrak{g})\mathfrak{g}$, so that (in X_σ at least)

$$\mathcal{D}_X \otimes_{U(\mathfrak{g})} V = \mathcal{D}_X \otimes_{U(\mathfrak{g})} U(\mathfrak{g})/U(\mathfrak{g})\mathfrak{g} = \mathcal{D}/\mathcal{D}\partial_u.$$

This is just $\mathcal{O}_X$. In fact in the other chart (with coordinate $x = 1/u$) we have

$$E \mapsto x^2\frac{\partial}{\partial x}$$
$$H \mapsto 2x\frac{\partial}{\partial x}$$
$$F \mapsto -\frac{\partial}{\partial x},$$

and we similarly conclude $\mathcal{V} = \mathcal{O}$. Just to labour the point a little further: in the first chart U we have $\mathcal{O}(U) \cong \mathcal{D}(U)/(\partial_u)$ and in the second open U' we have $\mathcal{O}(U') \cong \mathcal{D}(U')/(\partial_x)$. The restrictions to $U \cap U'$ go to $D = \mathbb{C}[u, u^{-1}, \partial_u]$ and in this ring we have an equality of left ideals

$$D\partial_u = D\partial_x.$$

EXAMPLE OF TRIVIAL MODULE BECOMING $\mathcal{O}_X$? EXAMPLE OF A NONTRIVIAL IRREP GOING TO ZERO.... ? THEN DEDUCE SOME CONDITION ON THE

CENTRAL CHARACTER. THEN TALK ABOUT LOCALISATION TO TWISTED DIF-
FERENTIAL OPERATORS.

For most $\mathfrak{g}$-modules M we have $\Delta(M) = 0$ in fact. This becomes clear upon considering the Casimir operator Ω. Indeed recall that

$$\Omega = EF + FE + \frac{1}{2}H^2$$

for $\mathfrak{g} = sl_2$. The image Ω' of this in $\mathcal{D}_X(U)$, where $U = X_\sigma$ is the big cell, is (using (**??**))

$$\Omega' = -\frac{\partial}{\partial u}u^2\frac{\partial}{\partial u} - u^2\frac{\partial}{\partial u}\frac{\partial}{\partial u} + 2u\frac{\partial}{\partial u}u\frac{\partial}{\partial u} = 0.$$

One finds similarly that the image of Ω is zero in the other open affine. Now let M be a $\mathfrak{g}$-module in which $\Omega = cI_M$ (for example any irreducible or Verma). If $c \neq 0$ then, of course,

$$P \otimes m = \frac{1}{c}(P \otimes cm) = \frac{1}{c}(P \otimes \Omega m) = \frac{1}{c}(P\Omega' \otimes m) = 0\frac{1}{c}(P \otimes m) = 0.$$

If we repeat this calculation to find the image of Ω in $\mathcal{D}_X^E$ for $E = \mathcal{L}_{-k\varpi}$ using (**??**), we find that Ω' equals some definite constant, which depends on k. Indeed what is supposed to happen is that the space of global sections

$$\Gamma(X, \mathcal{L}_{-\Lambda}) \cong L(\Lambda)^* \qquad (\text{for } \Lambda \in P_+)$$

becomes a $\mathcal{D}_X^E$-module. The Beilinson-Berstein localisations of Verma modules $M(\lambda)$ and their simple quotients $L(\lambda)$, are zero except possibly for λ in the same weak block as Λ.

———————————-

A straightforward (but by no means obvious) theorem:

$$\Gamma(G/B, \mathcal{D}_{-2\rho}) \cong U(\mathfrak{g})/(\chi(x) + 2\rho(x))$$

Uses the nilpotent cone and such...?

A really hard, and vitally important, theorem:

$$\mathcal{D}_{-2\rho} - mod \cong \Gamma(G/B, \mathcal{D}_{-2\rho}) - mod$$

There are $|W|$ Vermas in this category, and they correspond to $\mathcal{D}$-modules supported on cells in the Bruhat decomposition.

The irreducibles are harder to describe, they are what is called "minimal extensions" or "intermediate extensions".

After applying the Riemann-Hilbert correspondence, each Verma module corresponds to a constructible sheaf which is (cohomologically) supposed in a single degree and (sheaf-theoretically) supported on a single Bruhat cell.

The irreducible modules correspond to much more complicated things called IC-sheaves, which are spread out over several different cohomological degrees.
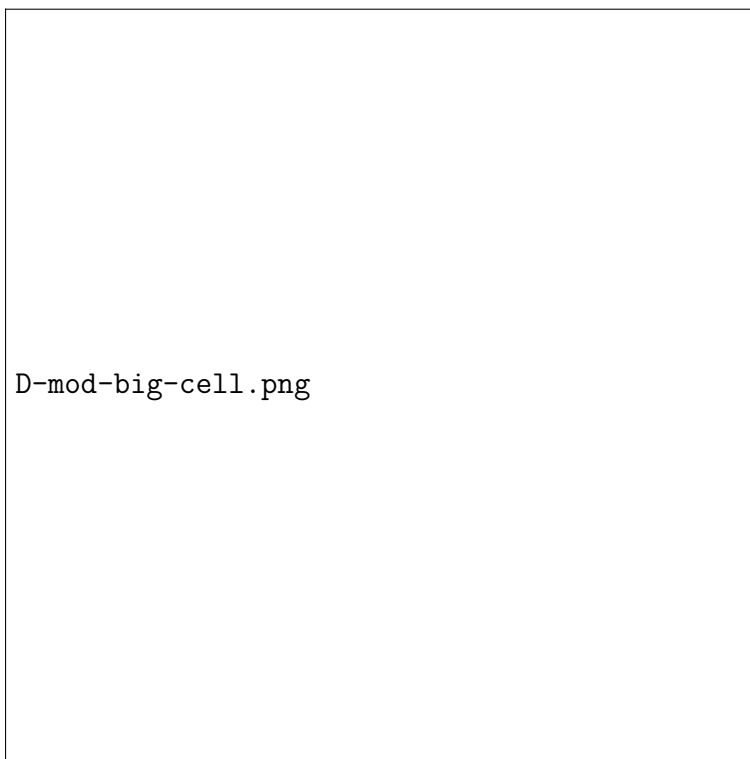
———————

OK great, so for each $w \in W$ we denote by $i_w : X_w \to X$ the inclusion, and we consider the $\mathcal{D}_X$-modules:

$$\mathcal{M}'_w = \int_{i_w} \mathcal{O}_{X_w}$$

First let's consider $w = \sigma$. We have $X_\sigma \cong \mathbb{A}^1$ open in X. So $\mathcal{M}_\sigma = \mathcal{O}_X(X_\sigma) \cong \mathbb{C}[u]$. If we draw the vectors u^k in a chain, it becomes easy enough to see that

$$\Gamma(X, \mathcal{M}'_\sigma) \cong M(0)^\vee$$

the dual Verma module!

Second let's consider $w = e$. Since the cell is $\{\infty\}$ let's work with the local coordinate $x = 1/u$. Relative to this coordinate we have $X_e = \{0\}$, and ∂_x is a normal vector field. We have $\mathcal{O}_* = \mathbb{C}$, let δ denote the image of 1 under the pushforward. Then the pushforward is spanned by δ, $\partial_x \delta$, $\partial_x^2 \delta$, etc. The action of x is determined by the relation $x\delta = 0$. For example this determines that

$$x(\partial \delta) = \partial(x\delta) - \partial(x)\delta = -\delta$$

So in fact

$$\Gamma(X, \mathcal{M}'_e) \cong M(-2)^{\vee}.$$

Remark 26.1. *A bit of a paradox: probably due to a sign error somewhere. Let us consider $\Delta(M(-2))$ and try to confirm the description we just obtained. We present $M(-2)$ as the quotient $U(\mathfrak{g})/U(\mathfrak{g}) < E, H + 2 >$ (quotient of left $U(\mathfrak{g})$-modules). In the big cell with coordinate u we find*

$$\mathcal{D}/\mathcal{D} < \partial_u, -2u\partial_u + 2 >= \mathbb{C}[u]/(-2u\partial_u + 2)\mathbb{C}[u]$$

It turns out that D-module dual corresponds to contragredient dual in category $\mathcal{O}$. So if we define $\mathcal{M}_w = \Delta(\mathcal{M}'_w)$ then $\mathcal{M}_e$ corresponds to $M(-2)$ and $\mathcal{M}_\sigma$ corresponds to $M(0)$.

———————

Next we need to explain

$$\mathcal{L}_e = \mathcal{M}_e,$$
$$\mathcal{L}_\sigma = \mathcal{M}_\sigma - \mathcal{M}_e.$$

So far there is nothing to explain as the LHS have not been defined. The point is that one defines the LHS in some way (specifically as IC sheaves, which means they are Verdier-self-dual) and a very deep theorem guarantees that they are irreducible in a certain sense (so their global sections correspond to the irreducible $\mathfrak{g}$-modules L_w).

113

This requires passing to perverse sheaves via the Riemann-Hilbert correspondence.

The other ingredient is to somehow associate these objects (holomorphic regular $\mathcal{D}_X$-modules) with elements of the Hecke algebra. In such a way (check details) that $\mathcal{M}_w$ corresponds to the basis element T_w, and $\mathcal{L}_w = \mathcal{M}_w + (lower order)$ has the property that it is invariant under the involution. This ends up proving the Kazhdan-Lusztig conjecture.

The annoying thing is that this requires even more weird theory: Hodge modules, which are a sort of refined hybrid of $\mathcal{D}$-modules and perverse sheaves.

————————

The explicit statements about where the Vermas and irreducibles go are as follows:

Theorem 26.2. *Let $i : X_w \to G/B$ denote the embedding of a Bruhat cell. There is an isomorphism of $\mathcal{D}$-modules*

$$\Delta(M_w^\vee) \cong \int_i \mathcal{O}_{X_w}.$$

Theorem 26.3 ([**?**, Proposition 12.3.2 (i)]). *Let $i : X_w \to G/B$ denote the embedding of a Bruhat cell. There is an isomorphism of $\mathcal{D}$-modules*

$$\Delta(L_w) \cong i_{!*}\mathcal{O}_{X_w}.$$

First we explain the notation. In general if $f : X \to Y$ and $\mathcal{M}$ is a $\mathcal{D}_X$-module, then there exists a canonical morphism

$$\int_{f!} \mathcal{M} \to \int_f \mathcal{M}$$

and we denote its image by $f_{!*}\mathcal{M}$. This $\mathcal{D}_Y$-module is called the "minimal extension" of $\mathcal{M}$ relative to f.

We fix notation as follows: $U = X_\sigma = \mathbb{P}^1 \backslash \{\infty\}$ and $V = \mathbb{P}^1 \backslash \{0\}$, with inclusions

$$j : U \to \mathbb{P}^1 \leftarrow \{\infty\} : i.$$

First compute $\int_j \mathcal{O}_U$ using the more standard triangle.

If $\mathcal{M}$ is a left $\mathcal{D}$-module, then we may consider

$$\operatorname{Hom}_{\mathcal{D}_X}(\mathcal{M}, \mathcal{D}_X).$$

We prefer to consider this operation only for relatively "small" $\mathcal{D}$-modules – we should like them to be coherent as $\mathcal{O}_X$-modules at least. Intuitively it is like taking the dual of a vector space, if the dimension is infinite then the dimension of the dual is very much larger. The result of this operation is now a right $\mathcal{D}_X$-module, so we convert it back (and throw in a shift for good measure) defining

$$\mathbb{D}\mathcal{M} = \operatorname{Hom}_{\mathcal{D}_X}(\mathcal{M}, \mathcal{D}_X) \otimes_{\mathcal{O}_X} \omega_X^{\otimes -1}[\dim_X].$$

Sheaves	$\mathcal{D}$-modules
f_*	$\int_f$
f^{-1}	$f^\star = \mathbb{D}_X \circ f^\dagger \circ \mathbb{D}_Y$
$f_!$	$\int_{f!} = \mathbb{D}_Y \circ \int_f \circ \mathbb{D}_X$
$f^!$	$f^\dagger = f^*[\dim_X - \dim_Y]$
$\mathbb{D}$	$\mathbb{D}$

Table 1: Matching between image functors for plain sheaves and $\mathcal{D}$-modules

This operation sends the category (bounded complexes of) coherent $\mathcal{D}$-modules into itself, and $\mathbb{D}^2 = \mathrm{Id}$ [**?**, Proposition 2.6.5].

The following is [**?**, Theorem 1.7.1].

Theorem 26.4. *Let $Z \subset X$ be a closed subset, assume Z is smooth as a variety. Let $i : Z \to X$ denote the inclusion, and $j : U \to X$ the inclusion of the complement. Then there exists a distinguished triangle*

$$\int_i i^\dagger \mathcal{M} \to \mathcal{M} \to \int_j j^\dagger \mathcal{M} \xrightarrow{+1}$$

If we use plain sheaf notation of Table **??**, then the distinguished triangle of the theorem above becomes

$$i_! i^! \mathcal{M} \to \mathcal{M} \to j_* j^* \mathcal{M} \xrightarrow{+1} . \tag{26.3}$$

According to [**?**, p. 19] and to [**?**, p. 18], there is another distinguished triangle:

$$j_! j^! \mathcal{M} \to \mathcal{M} \to i_* i^* \mathcal{M} \xrightarrow{+1}$$

referred to as the triangle "Verdier dual" to (**??**). Let's derive it. According to [**?**, p. 18] This should be somewhere in HTT, just have to find it. we have $\int_i = \int_{i!}$ for a closed inclusion, and $j^\star = j^\dagger$ for an open inclusion.

Opening up the definitions here gives

$$\mathbb{D}_X i_* \mathbb{D}_Z i^! \mathcal{M} \to \mathcal{M} \to j_* \mathbb{D}_U j^! \mathbb{D}_X \mathcal{M} \xrightarrow{+1}$$

Now let's apply this triangle not to $\mathcal{M}$ but to $\mathbb{D}(\mathcal{M})$, and let's also post-compose everything with $\mathbb{D}$. The post-composition reverses arrows. We get

$$\xleftarrow{+1} i_* \mathbb{D}_Z i^! \mathbb{D}_X \mathcal{M} \leftarrow \mathbb{D}_X^2 \mathcal{M} \leftarrow \mathbb{D}_X j_* \mathbb{D}_U j^! \mathcal{M}.$$

A potentially confusing detail: the left-most arrow still has degree $+1$ because the operation $\mathrm{Hom}_{\mathcal{D}}(-, \mathcal{D})$ is understood to send the term in degree $+k$ to degree $-k$ in general. Using the definitions of i_* and $j^!$ again, and $\mathbb{D}_X^2 = \mathrm{Id}$, we have

$$j_! j^! \mathcal{M} \to \mathcal{M} \to i_* i^* \mathcal{M} \xrightarrow{+1} .$$

as required.

Anyway, to compute $\int_j \mathcal{O}_U$ we use the standard triangle

$$\int_i i^\dagger \mathcal{M} \to \mathcal{M} \to \int_j j^\dagger \mathcal{M} \xrightarrow{+1}$$

with $\mathcal{M} = \mathcal{O}_{\mathbb{P}^1}$. The rightmost term is precisely $\int_j \mathcal{O}_U$. Now $i^\dagger \mathcal{O}_{\mathbb{P}^1}$ is the $\mathcal{O}$-module pullback with a homological shift. More precisely it is $\underline{\mathbb{C}}[-1]$ (i.e., concentrated in homological degree 1). On U nothing much is happening: we have $\int_j \mathcal{O}_U(U) = \mathcal{O}_{\mathbb{P}^1}(U)$. On V things are more interesting. Thus the LES is telling us

$$
\begin{array}{ccc}
\mathbb{C}[\partial_x] & \longrightarrow & 0 \\
& & \\
0 \longrightarrow & \mathcal{D}/\mathcal{D}\cdot\partial_x & \longrightarrow ?
\end{array}
$$

From this we infer

$$[\int_j \mathcal{O}_U](V) = \mathcal{D}/\mathcal{D}\cdot\partial_x x. \tag{26.4}$$

To compute $\int_{j!} \mathcal{O}_U$ we will use the triangle

$$\int_{j!} j^! \mathcal{M} \to \mathcal{M} \to \int_i i^* \mathcal{M} \xrightarrow{+1}$$

We take $\mathcal{M} = \mathcal{O}_{\mathbb{P}^1}$. It is a fact that

$$j^! \mathcal{O}_{\mathbb{P}^1} = j^* \mathcal{O}_{\mathbb{P}^1} = \mathcal{O}_U,$$

so the leftmost term is just what we are computing. Now we proceed by computing on U and on V. Firstly on U nothing much happens, we have

$$[\int_{j!} \mathcal{O}_U](U) = \mathcal{O}_{\mathbb{P}^1}(U) = \mathbb{C}[u].$$

On V things are a little more interesting. We claim that

$$[\int_{j!} \mathcal{O}_U](V) = \mathcal{D}/\mathcal{D}\cdot x\partial_x \tag{26.5}$$

(where $D = \mathbb{C}[x, \partial_x]$). First we note that $\mathcal{M}(V) = \mathcal{O}_{\mathbb{P}^1}(V) = \mathcal{D}/\mathcal{D}\cdot\partial_x$. Next we recall that $i^*\mathcal{M}$ is the $\mathcal{O}$-module pullback as a sheaf, so $i^*\mathcal{M} = \mathbb{C}$ on the point. Now a confusing point: the pullback $i^\dagger\mathcal{O}$ would be $\mathbb{C}[-1]$ because of the definition, i.e., concentrated in degree $+1$. But I think the triangle I am using is Verdier dual to the usual one, and i^* is supposed to

116

mean Verdier dual to $i^\dagger$, which in this case just moves $\mathbb{C}$ to degree -1. Now the pushforward is polynomials in ∂_x, concentrated in homological degree -1. Staring at the LES now
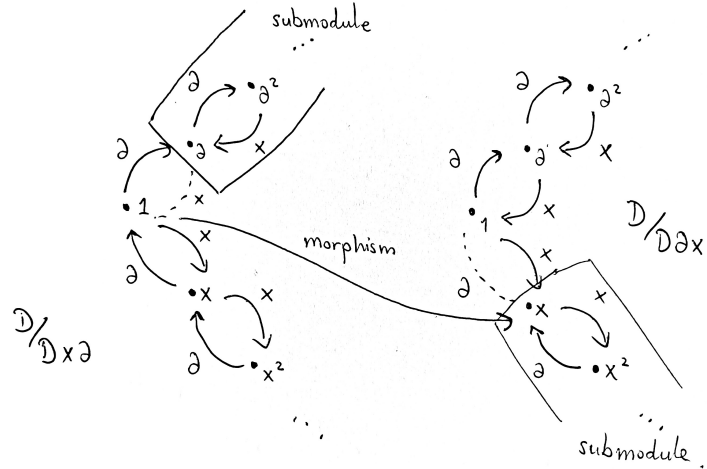
$$? \longleftarrow \mathcal{D}/\mathcal{D} \cdot \partial_x \longrightarrow 0$$

$$0 \longrightarrow \mathbb{C}[\partial_x]$$

we infer (??). The map in is just $\partial_x \mapsto \partial_x$ and the map out is the quotient.

So $\int_{j_!} \mathcal{O}_U \to \int_j \mathcal{O}_U$ is (1) on U it is the identity map, (2) on V it is the map

$$\mathcal{D}/\mathcal{D} \cdot x\partial_x \to \mathcal{D}/\mathcal{D} \cdot \partial_x x.$$

The picture is as follows

Inside $\mathcal{D}/\mathcal{D} \cdot x\partial_x$ we have the submodule generated by ∂_x, the quotient by this submodule is generated by 1. There is a well-defined morphism of $\mathcal{D}$-modules killing the submodule and sending $1 \mapsto x$. The image is isomorphic to $\mathbb{C}[x]$ with x^{k+1} playing the role of x^k. This might seem peculiar, but consider $\partial_x x = 0$ in this module, so indeed x is behaving like 1.

Anyway we have computed $j_{!*}\mathcal{O}_U$ to be $\mathbb{C}[u]$ on u and $\mathbb{C}[x]$ on V. One can check things match up as expected on the intersection, and indeed

$$j_{!*}\mathcal{O}_U \cong \mathcal{O}_{\mathbb{P}^1} = \Delta(L_\sigma).$$

27 Roadmap from here

The multiplicities of irreducibles L_y in Vermas M_w (here $w, y \in W$) are computed from a certain weird change-of-basis in the Hecke algebra. To make the connection, the idea roughly is to construct an associative product of $\mathcal{D}$-modules on $X = G/B$.

The way to do this is to consider the product variety $X \times X$, which carries two left actions of G – on the left and right factors, respectively. Denote by ΔG the diagonal action, i.e.,

$$g(x_1 B, x_2 B) = (g x_1 B, g x_2 B).$$

It is easy to check that as sets

$$X \times X / \Delta G \cong X/B = B \backslash G / B.$$

Indeed the action of G on X is transitive, so each ΔG-orbit contains an element of the form (eB, gB), and now we just note that

$$(eB, g_1 B) \sim (eB, g_2 B)$$

if and only if $g_2 = b g_1$ for some $b \in B$.

It follows that we can write

$$X \times X = \bigcup_{w \in W} Z_w$$

where the $Z_w = \Delta G(eB, n_w B)$ are ΔG-orbits, similarly to how we wrote

$$X = \bigcup_{w \in W} X_w$$

where the X_w are B-orbits.

It is now possible to introduce a product on D-modules. It is called the convolution product, and is actually an instance of quite a general construction. We consider the following pair of morphisms:

$$
\begin{array}{ccccc}
(a, b, c) & & X^3 & & (a, b, c) \\
 & \swarrow^{p_{13}} & & \searrow^{r} & \\
(a, c) & X^2 & & X^2 \times X^2 & (a, b, b, c) \\
\end{array}
$$

Now given $\mathcal{D}$-modules M and N on X^2, we define

$$[M] * [N] = \sum_{k \in \mathbb{Z}} (-1)^k H^k \int_{p_{13}} r^*(M \boxtimes N).$$

If M and N are ΔG-equivariant, then so is their product.

Then we consider coherent ΔG-equivariant $\mathcal{D}$-modules on $X \times X$. Call this category $\mathbb{H}$.

Turns out [?, Proposition 13.1.3] and $K(\mathbb{H}) \cong \mathbb{C}[W]$ with isomorphism

$$\left[\mathbb{D} \int_{i_w} \mathcal{O}_{Z_w} \right] \mapsto (-1)^{\ell(w)} T_w.$$

One finds that the irreducibles go to $(-1)^{\ell(w)}C_w|_{q=1}$. Now these elements lie in $\mathbb{C}[W]$ but their definition requires the passage to the Hecke algebra (they are characterised by being invariant under some involution).

So it is necessary to somehow put q into the structure of our category $\mathbb{H}$. Then it will turn out, roughly, that the involution corresponds to Verdier duality, and the invariance of the $\mathcal{L}_w$ comes from a deep result called the decomposition theorem.

Definition 27.1. Let $A = \mathbb{Z}[q^{1/2}, q^{-1/2}]$. The Hecke algebra H is quotient of the free associative A-algebra with generators T_s for $s = 1, \ldots, \ell$ by the relations

$$(T_i T_j)^{m_{ij}} = (T_j T_i)^{m_{ij}},$$

and

$$T_s^2 = (q-1)T_s + qT_e.$$

It turns out that H is a free A-module of rank equal to the cardinality $|W|$ of the Weyl group W, and a basis of H is given by

$$T_w = T_{s_{i_1}} T_{s_{i_2}} \cdots T_{s_{i_\ell}}, \qquad w \in W$$

where $w = s_{i_1} s_{i_2} \cdots s_{i_\ell}$ is a reduced expression (Remarkably, different reduced expressions for w yield the same element T_w).

There is an involution, denoted $X \mapsto \overline{X}$ on H (as a $\mathbb{Z}$-algebra) which interchanges $q^{1/2}$ with $q^{-1/1}$ and T_w with $(T_{w^{-1}})^{-1}$. There is a unique basis $\{C_w \mid w \in W\}$ of H with the property $\overline{C_w} = C_w$ for each $w \in W$, and

$$C_w = q^{-\ell(w)/2} \sum_{y \leq w} Q_{y,w}(q^{1/2}) T_y,$$

for some polynomials $Q_{y,w}(t)$ where $Q_{w,w}(t) = 1$ and $\deg(Q_{y,w}) \leq \ell(w) - \ell(y) - 1$. This is reasonably easy to prove, actually.

It turns out, furthermore, that $Q_{y,w}(q^{1/2}) = P_{y,w}(q)$ (i.e., only even powers of t appear in $Q_{y,w}(t)$).

Then the Kazhdan-Lusztig conjecture (now theorem) says:

Theorem 27.1. *At the level of characters*

$$[L_w] = \sum_{x \leq w} (-1)^{\ell(w) - \ell(x)} P_{x,w}(1)[M_x].$$

Here are some examples, which I have taken from the nice review paper [?]. For example for $G = SL_3$ the algebra H has a basis

$$T_e, \quad T_1, \quad T_2, \quad T_{12}, \quad T_{21}, \quad \text{and} \quad T_\circ,$$

119

and we have

$$
\begin{aligned}
C_e &= T_e \\
C_1 &= q^{-1/2}(T_1 + 1), \\
C_2 &= q^{-1/2}(T_2 + 1), \\
C_{12} &= q^{-1}(T_{12} + T_1 + T_2 + 1), \\
C_{21} &= q^{-1}(T_{21} + T_1 + T_2 + 1), \\
C_\circ &= q^{-3/2}(T_\circ + T_{12} + T_{21} + T_1 + T_2 + 1).
\end{aligned}
$$

Examining the expression for $C_\circ$, we see that the Weyl-Kac character formula is confirmed. In fact this particular expression is always particularly simple, i.e., $P_{x,w_\circ}(q) = 1$ for all x. For larger groups, the polynomials become more complicated. Already in the case of $G = SL_4$ one has

$$
\begin{aligned}
C_{2132} = q^{-2}(&T_{2132} + T_{213} + T_{212} + T_{232} + T_{132}+ \\
&+ T_{21} + T_{12} + T_{13} + T_{23} + T_1 + T_3 + (1+q)T_2 + (1+q)T_e).
\end{aligned}
$$

The proof revolves around the geometry of $X = G/B$, the Bruhat cells $X_w = BwB/B \subset X$ and the Schubert varieties $\overline{X}_w$. One has $X_x \subset \overline{X}_w$ if and only if $x \le w$.

Poincaré polynomial of $\overline{X}_w$:

$$
P_w(q) = \sum_{i \in \mathbb{Z}} \dim H^{2i}(\overline{X}_w) q^i
$$

The Kazhdan-Lusztig polynomials are Poincaré polynomials of intersection homology groups

$$
P_{x,w}(q) = \sum_{i \in \mathbb{Z}} \dim IH_x^{2i}(\overline{X}_w) q^i.
$$

There is also $IH(\overline{X}_w)$ (no subscript) around

$$
\sum_{i \in \mathbb{Z}} \dim IH^{2i}(\overline{X}_w) q^i = \sum_{x \le w} q^{\ell(x)} P_{x,w}(q).
$$

Sometimes Poincaré duality fails for H, but is always maintained by IH

(A) Let $D = D_X = \Gamma(X, \mathcal{D}_X)$, and let $U_0 = U(\mathfrak{g})/\ker(\chi_0)$. Prove that $D_X \cong U_0$.

(B) Let $\mathcal{C}'$ be the category of all D-modules, and $\mathcal{C}$ be the category of $\mathcal{O}$-quasicoherent $\mathcal{D}$-modules on X. There are functors of localisation and global sections. Highly nontrivial theorem: these are equivalences of categories.

There is the subcategory $\mathcal{C}'_0$ of $U(\mathfrak{b})$-finite finitely generated modules. This corresponds to $\mathcal{C}_0 \subset \mathcal{C}$ of $\mathcal{D}$-modules (1) killed by an ideal of finite codimension in $U(\mathfrak{b})$ and (2) with a $U(\mathfrak{b})$-stable good filtration.

Another important theorem: all objects of $\mathcal{C}_0$ are regular and holonomic.

(C) The Riemann-Hilbert correspondence [**?**, Theorem 7.2.2] says that if X is a smooth algebraic variety, then

$$DR : D^b_{rh}(D_X) \to D^b_c(X)$$

is a equivalence, and [**?**, Theorem 7.2.5] The image of $\mathcal{D}_X - mod_{rh}$ under DR is the abelian category of perverse sheaves.

By the way, the definition of DR_X appears on p. 103, and is basically

$$DR_X(M^\bullet) = \Omega_X \otimes^L_{D_X} M^\bullet.$$

This definition makes it look like one can work algebraically. But it's very important, actually, to first analytify. I have very little intuition for why it's so vital. But a remark is that the solution functor $\mathrm{Hom}_{D_X}(-, \mathcal{O}_X)$ is quite similar, in that it coincides with $\otimes$ up to shift or something. And it's quite clear that the solution functor gives drastically different results if we work analytically or algebraically, as the example the exponential function illustrates.

(The definition of the RHS involves the Bruhat stratification, it must enter the definition of the LHS too).

So Vermas M_w and simples L_w go to specific perverse sheaves. The Verma modules are each sent to a constructible sheaf supported on a single Bruhat cell. The simple objects are sent to "intersection cohomology" sheaves.

28   The Hecke algebra

The proof of the Borel-Weil theorem suggests that the set $B\backslash G/B$ is important. In fact functions on this set form a remarkable algebra called the Hecke algebra.

Let G be a finite group and B a subgroup of G. For the development of this course, the groups G and B will be the $\mathbb{F}_p$-points of a simple algebraic group G and B a Borel subgroup. But the general construction works in general.

If S is any set, we denote by $C(S)$ the vector space of functions $f : S \to \mathbb{C}$ with pointwise addition.

The convolution $* : C(G) \times C(G) \to C(G)$ is defined by

$$[f_1 * f_2](g) = \frac{1}{|G|} \sum_{x \in G} f_1(x) f_2(x^{-1}g).$$

Obviously $C(G/B)$ is naturally identified with the subset

$$\{f : G \to \mathbb{C} | f(gb) = f(g) \text{ for all } g \in G, \, b \in B\},$$

of $C(G)$. Then

$$[f_1 * f_2](gb) = \frac{1}{|G|} \sum_{x \in G} f_1(x) f_2(x^{-1} gb),$$

and so $f_1 * f_2 \in C(G/B)$ whenever $f_2 \in C(G/B)$. So we have

$$* : C(G) \times C(G/B) \to C(G/B).$$

In a similar way we have

$$* : C(B\backslash G) \times C(G) \to C(B\backslash G).$$

The second one is the same argument, but with a change of summation variable in the middle.

The Hecke algebra associated to (G, B) is now defined to be $H(G, B) := C(B\backslash G/B)$ with the convolution product, which it is indeed closed under.

If G is the $\mathbb{F}_q$-points of a simple/semisimple/reductive algebraic group and $B \subseteq G$ the corresponding Borel then there is a theorem that $B\backslash G/B$ is in bijection with W the Weyl group (Bruhat decomposition?).

This Hecke algebra can be explicitly described (first studied by Iwahori?) by generators (the T_s associated to generating reflections $s \in W$) and relations, which are like

$$(T_s + 1)(T_s + q) = 0.$$

Comparing with the usual answer for $SL_2(\mathbb{Z}_2)$ checks out except the normalisation is a bit off. Probably the standard normalisation is to divide by $|B|$ rather than $|G|$.

More symmetrically

$$[f_1 * f_2](g) = \sum_{xy=g} f_1(x) f_2(y).$$

The double coset of g only depends on Bx and yB. So we can figure out the convolution of two indicator functions (on the double cosets $Bw_1 B$ and $Bw_2 B$ say) by writing $Bw_1 B = \cup_i Ba_i^{(1)}$ and $Bw_2 B = \cup_j a_j^{(2)} B$. The we look at the products $a_i^{(1)} a_j^{(2)}$ and count how many lie in each double coset, these numbers give the coefficients of the indicator functions that occur in the product of our two double cosets.

Though if you think about it, the convolution $\delta_B * \delta_B$ is now $|B|\delta_B$ because each $x \in B$ contributes 1 in the sum and they all contribute to the double coset B. So indeed the correct normalisation is to divide by $|B|$.

29 The (weak) BGG resolution

Let $\mathfrak{g}$ be a finite dimensional simple Lie algebra. Recall the Weyl character formula

$$\mathrm{Ch}_{L(\Lambda)} = \sum_{w \in W} \varepsilon(w) \, \mathrm{Ch}_{M(w \circ \Lambda)},$$

valid whenever $\Lambda \in P_+$. Imagine there were a complex $D_\bullet$ in which $D_0 = M(\Lambda)$, $D_1 = \bigoplus_{i=1}^\ell M(s_i \circ \Lambda)$ and in general $D_n = \bigoplus_{\ell(w)=n} M(w \circ \Lambda)$, so that

$$H_i(D_\bullet) = \begin{cases} L(\Lambda) & i = 0 \\ 0 & i \neq 0 \end{cases}$$

That would be very nice since, because $\varepsilon(w) = (-1)^{ell(w)}$, the Poincaré principle applied to $D_\bullet$ would immediately yield the Weyl character formula.

In fact it is true, and the resolution is called the BGG resolution. We will construct something a little weaker than $D_\bullet$ as described above, instead we will construct a complex $D_\bullet$ for which each D_n carries a filtration

$$D_n = D_n^{(0)} \supset D_n^{(1)} \supset \ldots \supset D_n^{(N)} = 0,$$

whose successive quotients are precisely the $M(w \circ \Lambda)$ for $\ell(w) = n$ in some order. Since a filtered module and its associated graded always have the same character, the existence of this weak BGG resolution is still sufficient to prove the Weyl character formula.

The material in this section is taken quite directly from [?, Sections 6.1-6.4].

––––––––––––

We will construct a first approximation to the desired complex by imitating the Chevalley-Eilenberg complex of usual Lie algebra homology. Instead of

$$V_n = U(\mathfrak{g}) \otimes_{\mathbb{C}} \wedge^n \mathfrak{g},$$

we take

$$D_n = U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} \wedge^n(\mathfrak{g}/\mathfrak{b}).$$

The differential looks similar, in detail it will be

$$d(u \otimes \xi_1 \wedge \ldots \wedge \xi_n) = \sum_{i=1}^n (-1)^{i+1}(uz_i) \otimes \xi_1 \wedge \ldots \wedge \widehat{\xi_i} \wedge \ldots \wedge \xi_n$$
$$+ \sum_{i<j} (-1)^{i+j} u \otimes \overline{[\xi_i, \xi_j]} \wedge \xi_1 \wedge \ldots \wedge \widehat{\xi_i} \wedge \ldots \wedge \widehat{\xi_j} \wedge \ldots \wedge \xi_n.$$

Here $\xi_i \in \mathfrak{g}/\mathfrak{b}$ and z_i means any representative in $\mathfrak{g}$ of ξ_i. Then the overline means to project back to $\mathfrak{g}/\mathfrak{b}$. In addition to checking that $d \circ d = 0$ we have to prove that d is actually well-defined. For this it is essential that the tensor product is over $U(\mathfrak{b})$, which we use to pass elements of $\mathfrak{b}$ back and forth across it.

Proving that $D_\bullet$ is exact, in general, requires a spectral sequence argument like in the usual case. If we happen to be in a situation where $\mathfrak{g} = \mathfrak{b} \oplus \mathfrak{n}_-$ a splitting of vector spaces in which $\mathfrak{b}$ and $\mathfrak{n}_-$ are both subalgebras... then we can cheat and prove the result very easily.

The idea is this: by the PBW theorem we have

$$U(\mathfrak{g}) \cong U(\mathfrak{n}_-) \otimes_{\mathbb{C}} U(\mathfrak{b}),$$

we also have $\mathfrak{g}/\mathfrak{b} \cong \mathfrak{n}_-$. So we have vector space isomorphisms

$$U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} \wedge^n(\mathfrak{g}/\mathfrak{b}) \cong U(\mathfrak{n}_-) \otimes_{\mathbb{C}} \wedge^n(\mathfrak{n}_-),$$

choose all coset representatives to actually lie in $\mathfrak{n}$. Then the differential coincides exactly with the usual Chevalley-Eilenberg differential for $\mathfrak{n}_-$. So we have an isomorphism of complexes at the level of vector spaces, which yields exactness.

Remark 29.1. *One might ask, given that $D_\bullet(\mathfrak{g}, \mathfrak{b}) \cong V_\bullet(\mathfrak{n}_-)$, what advantage is gained by introducing $D_\bullet(\mathfrak{g}, \mathfrak{b})$? One way to say it would be to observe that $V_\bullet(\mathfrak{n}_-)$ acquires a $\mathfrak{g}$-module structure compatible with the differentials. If we did not know the isomorphism to $D_\bullet(\mathfrak{g}, \mathfrak{b})$ then this would not be at all easy to see. Indeed the action of $x \in \mathfrak{g}$ would involve multiplying by x from the left, commuting it through, writing it as $x = n + b$, letting b pass to the exterior algebra factors.*

––––––––––––

Sumarising, we have

$$H_i(D_\bullet) = \begin{cases} \mathbb{C} & i = 0 \\ 0 & i \neq 0 \end{cases}$$

a resolution of the trivial $\mathfrak{g}$-module $\mathbb{C}$. We would like to check that the D_n are filtered by Verma modules. This follows from the general Lemma **??**.

This lemma also tells which Verma factors appear in each D_n and with what multiplicities. A little thought reveals that D_n is much larger than $\{w \in W \mid \ell(w) = n\}$. Indeed if that were the case, we would have a total of $\#W$ Vermas, but actually if $m = \dim(\mathfrak{g}/\mathfrak{b})$ we have $\dim \wedge^\bullet(\mathfrak{g}/\mathfrak{b}) = 2^m$ and so we have 2^m Vermas, which is much more.

Luckily $D_\bullet$ splits into many subcomplexes, only one of which has a nonzero morphism to $\mathbb{C}$, we need to single out this summand, and we will see that it has the desired number of Verma factors.

Next we are going to prove a key property of the weights occuring in the $\mathfrak{b}$-module $\wedge^*(\mathfrak{g}/\mathfrak{b})$. First we need some lemmas.

Definition 29.1. For a subset $\Sigma \subset \Delta_+$ we will denote by $|\Sigma|$ the sum of all the elements in Σ.

So for example $|\Delta_+| = 2\rho$.

Definition 29.2. For $w \in W$ we will denote

$$\Pi_w = \Delta_+ \cap w(\Delta_-).$$

The first lemma is this:

Lemma 29.2.
$$w \circ 0 = -|\Pi_w|.$$

We will also require two other lemmas which we have already proved. Namely

Lemma 29.3. *For each simple root α_i we have*

$$s_i(\Delta_+ \backslash \{\alpha_i\}) = \Delta_+ \backslash \{\alpha_i\}.$$

And

Lemma 29.4. *Let $w \in W$ of length $\ell(w) = k$. Then there exists i such that $r_i w$ has length $k - 1$. Furthermore*

$$w^{-1}(\alpha_i) \in \Delta_- \qquad and \qquad (r_i w)^{-1}(\alpha_i) \in \Delta_+.$$

(Although not identical, convince yourself that this follows from Lemma **??**.)

Proposition 29.5. *For each $w \in W$ the weight $w \circ 0$ appears among the weights of $\wedge^*(\mathfrak{g}/\mathfrak{b})$ with multiplicity 1.*

Proof. We write $m = \dim(\mathfrak{g}/\mathfrak{b}) = \#\Delta_+$. The dimension of $\wedge^*(\mathfrak{g}/\mathfrak{b})$ is 2^m with a contribution of 1 from each subset $\Sigma \subset \Delta_+$. The nonzero basis vector contributed by Σ is

$$\pm e_{-\alpha_1} \wedge \cdots \wedge e_{-\alpha_n} \quad \text{where} \quad \Sigma = \{\alpha_1, \ldots \alpha_n\}.$$

Suppose $\mu = w \circ 0$, then by Lemma **??** we have $\mu = -|\Pi_w|$. This interpretation allows us to deduce the multiplicity 1 property from the following lemma. $\qquad \square$

Lemma 29.6. *Let $w \in W$ and $\Sigma \subset \Delta_+$. If $|\Sigma| = |\Pi_w|$ then $\Sigma = \Pi_w$.*

Proof. The proof is by induction on the length $\ell(w)$. For the base case we note that $\ell(w) = 0$ only for $w = e$, we have $\Pi_e = \emptyset$ and $|\Pi_e| = 0$, and then since all the elements of Δ_+ are on one side of a hyperplane, clearly $|\Sigma| = 0$ only for $\Sigma = \emptyset = \Pi_e$.

Now we suppose the result has been established for all w of length at most k, and choose w of length $\ell(w) = k > 0$. There exists a simple root $\alpha \in \Pi$ such that

$$\ell(s_\alpha w) = k - 1$$
$$w^{-1}(\alpha) \in \Delta_-$$
$$(s_\alpha w)^{-1}(\alpha) \in \Delta_+.$$

Write $w' = s_\alpha w$. In particular $\alpha \in \Pi_w$ and $\alpha \notin \Pi_{w'}$.

Recall that

$$s_\alpha(\Delta_+ \backslash \{\alpha\}) = \Delta_+ \backslash \{\alpha\}.$$

From this we will deduce that

$$\Pi_w = s_\alpha(\Pi_{w'}) \cup \{\alpha\}.$$

Indeed

$$\begin{aligned}
s_\alpha(\Pi_{w'}) &= s_\alpha(\Delta_+) \cap s_\alpha(w'(\Delta_-)) \\
&= ((\Delta_+ \backslash \{\alpha\}) \cup \{-\alpha\}) \cap w(\Delta_-) \\
&= (\Delta_+ \backslash \{\alpha\}) \cap w(\Delta_-) \qquad (\text{since } \{-\alpha\} \cap w(\Delta_-) = \emptyset) \\
&= \Pi_w \backslash \{\alpha\}.
\end{aligned}$$

Now suppose that $\Sigma \subset \Delta_+$ with $|\Sigma| = |\Pi_w|$. Since we have $\alpha \in \Pi_w$ and $\alpha \notin \Pi_{w'}$, and we are supposed to be proceding by induction, it seems like a good idea to try to relate $\Pi_{w'}$ with $\Sigma \backslash \{\alpha\}$.

First suppose $\alpha \in \Sigma$. Then

$$|\Sigma \backslash \{\alpha\}| = |\Sigma| - \alpha = |\Pi_w| - \alpha = \rho - w(\rho) - \alpha,$$

whence

$$|s_\alpha(\Sigma \backslash \{\alpha\})| = s_\alpha |\Sigma \backslash \{\alpha\}| = s_\alpha(\rho - w(\rho) - \alpha) = \rho - \alpha - w'(\rho) + \alpha = \rho - w'(\rho) = |\Pi_{w'}|.$$

Hence by induction $s_\alpha(\Sigma \backslash \{\alpha\}) = \Pi_{w'}$, and so

$$\Sigma = s_\alpha(\Pi_{w'}) \cup \{\alpha\} = \Pi_w$$

as required.

Next suppose $\alpha \notin \Sigma$. We shall derive a contradiction, thus completing the proof. Well in this case $\Sigma \subset \Delta_+ \backslash \{\alpha\}$, so $s_\alpha(\Sigma) \subset \Delta_+ \backslash \{\alpha\}$ and thus

$$s_\alpha(\Sigma) \cup \{\alpha\} \subset \Delta_+.$$

By a similar calculation as above, we have

$$\begin{aligned}
|s_\alpha(\Sigma) \cup \{\alpha\}| &= s_\alpha |\Sigma| + \alpha = s_\alpha |\Pi_w| = s_\alpha(\rho - w(\rho)) + \alpha \\
&= \rho - \alpha - w'(\rho) + \alpha = \rho - w'(\rho) = |\Pi_{w'}|.
\end{aligned}$$

By induction $s_\alpha(\Sigma) \cup \{\alpha\} = \Pi_{w'}$, but we already know that $\alpha \notin \Pi_{w'}$, so there is a contradiction and this case (namely $\alpha \notin \Sigma$) does not occur.

$\square$

OK that was pretty exhausting, but we have the conclusion that

30   The Chevalley restriction theorem

The material of this section is explained in [**?**, Chapter 10].

Let V be a vector space over $\mathbb{C}$. We denote by $\mathbb{C}[V]$ the algebra of polynomial functions on V. (If we choose a basis of the dual of V then each of these is by definition a linear, hence polynomial, function on V and $\mathbb{C}[V]$ is the unital algebra generated by these.)

Let K be a group and V a K-module. Then we have a representation of K on $\mathbb{C}[V]$ by

$$[g \cdot f](v) = f(g^{-1} \cdot v).$$

We denote by $\mathbb{C}[V]^K$ the set of K-invariant elements, which is a subalgebra.

Suppose K is a finite group now. Then it is possible to show that $\mathbb{C}[V]^K$ is finitely generated.

Theorem 30.1 (Chevalley restriction theorem). *Let G be a simple complex Lie group with Lie algebra $\mathfrak{g}$ and Cartan subalgebra $\mathfrak{h}$. Then under the restriction map*

$$\rho : \mathbb{C}[\mathfrak{g}] \to \mathbb{C}[\mathfrak{h}],$$

we have

$$\rho(\mathbb{C}[\mathfrak{g}]^G) = \mathbb{C}[\mathfrak{h}]^W.$$

Proof. For definiteness let us write σ for the map $\mathbb{C}[\mathfrak{g}]^G \to \mathbb{C}[\mathfrak{h}]$. First we show that σ is injective.

Let $f \in \ker(\sigma)$, which is to say $f(h) = 0$ for all $h \in \mathfrak{h}$. At the same time f is G-invariant. So if we can show $\mathrm{Ad}(G)\mathfrak{h} \subset \mathfrak{g}$ is Zariski dense then we will know that $f = 0$. So consider the action

$$\varphi : G \times \mathfrak{h} \to \mathfrak{g}$$
$$(g, h) \mapsto \mathrm{Ad}(g)h$$

At a point $p = (e, h)$, the differential $d\varphi_p$ is

$$d\varphi_p(X, h') = h' + [X, h].$$

If we choose h so that $\alpha(h) \neq 0$ for all $\alpha \in \Delta$ then we see that $d\varphi_p$ is surjective. This implies that the image of φ contains an open set in the metric topology, and hence is Zariski dense in $\mathfrak{g}$.

Now we show that $\sigma(\mathbb{C}[\mathfrak{g}]^G) \subset \mathbb{C}[\mathfrak{h}]^W$. We think of W as

$$W = N_G(H)/H,$$

The action of w on $\mathfrak{h}$ coincides with the adjoint action $\mathrm{Ad}(\widetilde{w})$ of any chosen lift $\widetilde{w}$ of w to $N_G(H)$. But since the restriction of Ad-invariant functions are Ad-invariant, the inclusion follows immediately.

Finally we show equality $\sigma(\mathbb{C}[\mathfrak{g}]^G) = \mathbb{C}[\mathfrak{h}]^W$.

Recall $L(\lambda)$ the finite dimensional irreducible r$\mathfrak{g}$-module (and G-module by exponentiation) of highest weight $\lambda \in P_+$. Set $f_{\lambda,m} \in \mathbb{C}[\mathfrak{g}]$ defined by

$$f_{\lambda,m}(X) = \mathrm{Tr}_{L(\lambda)} X^m.$$

We claim that $f_{\lambda,m} \in \mathbb{C}[\mathfrak{g}]^G$, which is easy to see because

$$\mathrm{Tr}_{L(\lambda)} gXg^{-1} = \mathrm{Tr}_{L(\lambda)} X.$$

If we can show that $\sigma(f_{\lambda,m})$ span $\mathbb{C}[\mathfrak{h}]^W$ then we will be done.

Now we conisder also the *formal* character of $L(\lambda)$, namely

$$\chi_{L(\lambda)} = \sum_{\mu \in \mathfrak{h}^*} \dim(L(\lambda)_\mu) e^\mu \in \mathbb{C}[P],$$

where $\mathbb{C}[P]$ denotes the $\mathbb{C}$-algebra $\bigoplus_{\mu \in P} \mathbb{C} e^\mu$ with product $e^\mu \cdot e^{\mu'} = e^{\mu+\mu'}$. Alternatively think of $\mathbb{C}[P]$ as the collection of all functions $f : P \to \mathbb{C}$ with finite support.

There are two natural products on $\mathbb{C}[P]$ – considered as set functions $P \to \mathbb{C}$ – the pointwise product, which is what is implicit when we speak of the algebra $\mathbb{C}[V]$ above, and the convolution product.

The action of W on $\mathfrak{h}$ induces an action on $\mathfrak{h}^*$, under which P is invariant. So we can speak of $\mathbb{C}[P]^W$. We claim that $\mathbb{C}[P]^W$ is spanned by $\chi_{L(\lambda)}$ for $\lambda \in P_+$. Note first that each W-orbit in P meets P_+. Now let $f \in \mathbb{C}[P]^W$ and cosider its support, let $\mu \in P$ be the highest weight in the support of f. Proceed by reverse induction on height now, the weight μ lies in P_+, and we subtract

$$f - f_\mu \chi_{L(\mu)}$$

to get another W-invariant function of strictly lower height. That dispenses with the claim.

Next we consider

$$\mathbb{C}[P] \to \mathbb{C}[[\mathfrak{h}]] \to \mathbb{C}[\mathfrak{h}]_m$$

given by

$$e^\lambda \mapsto \sum \lambda^j / j!, \quad \text{and} \quad (\text{extract homogeneous degree } m \text{ component}).$$

We prove $\mathbb{C}[P]^W \to \mathbb{C}[\mathfrak{h}]_m^W$ is surjective, then we will be done. First note that $\mathbb{C}[P] \to \mathbb{C}[\mathfrak{h}]_m$ is surjective. Why is this? Well we could grab a basis $\{\varpi_1, \ldots, \varpi_\ell\}$ of P, on which the map gives $e^{\varpi_i} \mapsto \varpi_i^m$. These images do not span, but if we also take $e^{\varpi_1 + \varpi_2}$, for instance, then it goes to $(\varpi_1 + \varpi_2)^m$, and together we get all the degree m monomials as linear combinations. So grab $f \in \mathbb{C}[\mathfrak{h}]_m^W$, choose a preimage in $\widetilde{f} \in \mathbb{C}[P]$, do an averaging

$$\widehat{f} = \frac{1}{\#W} \sum_{w \in W} w\widetilde{f} \in \mathbb{C}[P]^W$$

and observe that $\widehat{f} \mapsto f$. $\qquad\qquad\square$

Proposition 30.2. *Let $X \in \mathfrak{g}$, then the intersection of $\mathrm{Ad}(G) \cdot X$ with $\mathfrak{h}$ is either empty or a single W-orbit.*

Proof. Since for every $w \in W$ the action of w on $\mathfrak{h}$ is given by the action of some $\widetilde{w} \in G$, the intersection (if nonempty) must be a union of W-orbits. Now let $h_1, h_2 \in \mathfrak{h}$ in the same G-orbit, we wish to show that they are in the same W-orbit.

We are supposed to use $\rho(\mathbb{C}[\mathfrak{g}]^G) = \mathbb{C}[\mathfrak{h}]^W$ somehow. To do so it helps to think of the points h_1, h_2 algebraically: they correspond to two homomorphisms ev_{h_1}, ev_{h_2} of $\mathbb{C}$-algebras $\mathbb{C}[\mathfrak{g}] \to \mathbb{C}$ (namely, evaluation of a function at the point). By composing with σ they give two homomorphisms $\mathbb{C}[\mathfrak{g}] \to \mathbb{C}$ (this just corresponds to the) $\qquad\square$

Let $Z(\mathfrak{g})$ denote the centre of $U(\mathfrak{g})$, i.e.,

$$Z(\mathfrak{g}) = \{z \in U(\mathfrak{g}) \mid zu = uz \text{ for all } u \in U(\mathfrak{g})\}.$$

Lemma 30.3. *Let $M(\Lambda)$ be a Verma module for $\mathfrak{g}$. Every $z \in Z(\mathfrak{g})$ acts in $M(\Lambda)$ as some constant.*

Exercise 30.4. *Prove the lemma.*

Let M be a $\mathfrak{g}$-module, and suppose each $z \in Z(\mathfrak{g})$ acts in M by a constant, which we denote $\chi(z)$. Then $\chi : Z(\mathfrak{g}) \to M$ is a homomorphism, which we call the central character of M.

The Harish-Chandra theorem.

Theorem 30.5. *Suppose $L(\lambda)$ is a composition factor of $M(\Lambda)$. Then $\lambda \leq \Lambda$ and $\lambda = w \circ \Lambda$ for some $w \in W$.*

Proof. From Lemma **??** we know that $M(\Lambda)$ has some central character $\chi = \chi_\Lambda$. Every submodule and quotient (and therefore $L(\lambda)$) has the same central character. That is,

$$\chi_\lambda = \chi_\Lambda.$$

$\qquad\square$

31 The Nilpotent cone

Theorem of Kostant is 3.2.5 in the book Chriss-Ginzburg.

This theorem is needed later in the proof that G/B is D-affine.

32 Pushforward with compact support

...

I see this used in the material on constructible sheaves, but not really for $\mathcal{D}$-modules. Indeed for $\mathcal{D}$-modules I see f^* and $f^!$, with the latter being defined as a homological shift of f^* (HTT p. 33 and Bernstein p. 5).

On the other Verdier duality is an essential ingredient of the perverse sheaves story, and it totally uses $f_!$ and its left adjoint. So I guess I need to cover this stuff.

Let $f : X \to Y$ be a continuous map of topological spaces. By definition the proper direct image $f_!\mathcal{F}$ is defined by

$$f_!\mathcal{F}(V) = \{s \in \mathcal{F}(f^{-1}(V)) \mid \text{The restriction } f|_{\text{supp}(s)} : \text{supp}(s) \to V \text{ is a proper map}\}$$

By its definition we have an injective morphism of sheaves $f_!\mathcal{F} \to f_*\mathcal{F}$.

The support of a section of a sheaf is the set of points at which the germ is nonzero. The support of a sheaf is a different notion, it is the set of points at which the stalk is nonzero. The support of a section of any sheaf is always a closed set. Not necessarily so for the support of a sheaf. [Hartshorne, II. Exercise 1.14, p. 67]

One can find the definition of "proper morphism" of schemes in Hartshorne, p. 100. But here we are talking about topological spaces in general. The definition is

Definition 32.1. A continuous function $f : X \to Y$ is proper if for all $K \subset Y$ compact, the imverse image $f^{-1}(K)$ is compact.

If f is a proper map then evidently the preimage of every point is compact. Perhaps this is a sensible intuition for the notion of proper.

We reserve the symbol $\Gamma_c(X, -)$ for $\pi_!(-)$ the compactly supported pushforward along the surjection $\pi : X \to *$ to a point.

Recovering Poincare from Verdier is HTT Example C.2.16

33 The derived category

This section based on [**?**, Chapter 2].

Let $\mathcal{A}$ be an abelian category. A complex of objects of $\mathcal{A}$ is a collection of objects and morphisms as follows

$$\cdots \to A^{-1} \to A^0 \to A^1 \to \cdots,$$

in which $d \circ d = 0$. A morphism $A \to B$ of complexes is a set of $f^i : A^i \to B^i$ such that all the obvious squares commute. There is a category, denoted $Kom(\mathcal{A})$, whose objects are complexes and morphisms are morphisms of complexes.

This category is again abelian (Proposition 2.3). It comes with an operation called "shift" or "suspension": $A \mapsto A[1]$ defined by $A[1]^i = A^{i+1}$, and $d[1]^i = -d^{i+1}$. Think of $A[1]$ as "A shifted one step left".

Given a complex A, one can form cohomology

$$H^i(A) = \frac{\ker(d : A^i \to A^{i+1})}{im(d : A^{i-1} \to A^i)},$$

thus obtaining an infinite sequence of elements of $\mathcal{A}$.

The notion of short exact sequence applies to any abelian category, and it turns out that

$$0 \to A \to B \to C \to 0$$

being a SES is equivalent to

$$0 \to A^i \to B^i \to C^i \to 0$$

being a SES in $\mathcal{A}$, for each i.

By the magic of diagram chasing, a SES of complexes, yields a LES of cohomology

$$\cdots \to H^i(A) \to H^i(B) \to H^i(C) \to H^{i+1}(A) \to \cdots,$$

This is the quintessential construction of homological algebra.

Homotopy and quasi-isomorphism.

Let f, g be two morphisms $A \to B$. They induce two morphisms $H^i(A) \to H^i(B)$, for each i, which we denote $H^i(f)$ and $H^i(g)$. We say f and g are "homotopic" if there exist $h^i : A^i \to B^{i-1}$ such that

$$f^i - g^i = d \circ h^i + h^{i+1} \circ d.$$

We also say "f is nullhomotopic" if f is homotopic to 0. Homotopy is an equivalence relation, and has the wonderful property that $H^i(f) = H^i(g)$ for homotopic maps f, g.

If we are given morphisms $f : A \to B$ and $g : B \to A$ whose compositions are homotopic to the identity morphisms on A and B, then they induce inverse isomorphisms of all homology objects. Each of f, g is then an example of a "quasi-isomorphism" (qis), that is, a morphism of complexes which induces isomorphisms between all corresponding cohomology objects. Such pairs (f, g) are called homotopy equivalences.

The homotopy category of $\mathcal{A}$, denoted $K(\mathcal{A})$, is defined to have the same set of objects as $Kom(\mathcal{A})$, and for objects A, B the set of morphisms is the quotient of the set of morphisms in $Kom(\mathcal{A})$ by the relation of homotopy. It's still additive, but no longer an abelian category. Many objects of $Kom(\mathcal{A})$ that were not isomorphic in that category, become isomorphic in $K(\mathcal{A})$. Indeed objects related by a pair of homotopy equivalences, become isomorphic.

The derived category of $\mathcal{A}$ appears when we attempt to make, not just homotopy equivalences, but all quasi-isomorphisms into isomorphisms between corresponding objects. The

derived category of $\mathcal{A}$, denoted $D(\mathcal{A})$, is defined to have the same set of objects as $Kom(\mathcal{A})$, and for objects A, B a morphism from A to B is an equivalence class of "roofs":

$$
\begin{array}{ccc}
 & C & \\
{\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot} \\
A & & B
\end{array}
$$
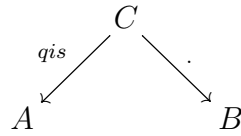
We will explain the equivalence relation between roofs later on, after describing composition of roofs.

First note that in particular every morphism $f : A \to B$ of complexes gives rise to a morphism between corresponding objects in $D(\mathcal{A})$:
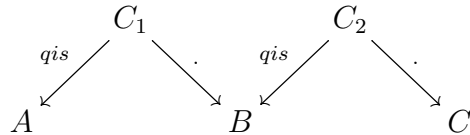
$$
\begin{array}{ccc}
 & A & \\
{\scriptstyle id}\swarrow & & \searrow{\scriptstyle f} \\
A & & B
\end{array}
$$
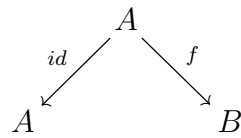
But one also has more curious morphisms like, if $q : B \to A$ is a quasi-isomorphism, then

$$
\begin{array}{ccc}
 & B & \\
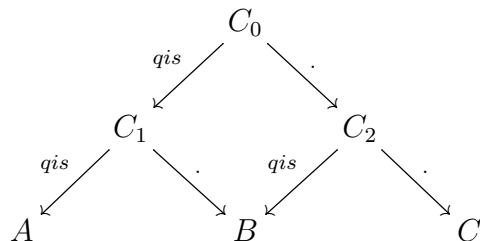{\scriptstyle q}\swarrow & & \searrow{\scriptstyle id} \\
A & & B
\end{array}
$$
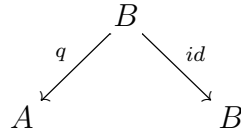
is a morphism $A \to B$ in $D(\mathcal{A})$, which functions as an inverse of q. Before checking that this is an inverse of q we need to define composition of morphisms in $D(\mathcal{A})$.

Let us consider two compatible morphisms, which we would like to compose

$$
\begin{array}{ccccc}
 & C_1 & & C_2 & \\
{\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot}\ {\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot} \\
A & & B & & C
\end{array}
$$

We would like to construct

$$
\begin{array}{ccccc}
 & & C_0 & & \\
 & {\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot} & \\
 & C_1 & & C_2 & \\
{\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot}\ {\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot} \\
A & & B & & C
\end{array}
$$

Such that the square commutes. (Why do we care if the square commutes or not?) If we ignored the condition that $C_0 \to C_1$ be a qis, there would be a natural thing to do: define $C_0 = C_1 \times_B C_2$, with its canonical projections.

Perhaps it is worth remarking that, given a category $\mathcal{C}$, there is something called the category of spans (or correspondences) of $\mathcal{C}$, whose objects are roofs without the qis condition. Then if $\mathcal{C}$ has fibred products, then the definition above is the composition in this category.

But with the qis condition in place, this does not work. We require an idea from algebraic topology (not the first, as we saw homotopy above already).

The cone of a morphism. Let $f : A \to B$ be a morphism of complexes. Then we define a new complex $C(f)$ as follows:

$$C(f)^i = A^{i+1} \oplus B^i, \qquad d(a,b) = (d_A(a), d_B(b) + f(a))$$

I like to draw the cone this way, even though it takes a lot of space.

Suppose we have two morphisms $f : A \to B$ and $g : B \to C$, so that we also have the composition $g \circ f$. Pondering the cones of these morphisms for a while, we see that there are canonical morphisms as shown:

Now let $f : A \to B$ again, and $C(f)$ the cone of this morphism. Then in fact we have a SES

$$0 \to B \to C(f) \to A[1] \to 0,$$

(denote the morphisms here τ and π) and thus a LES

$$\cdots \to H^i(B) \to H^i(C(f)) \to H^{i+1}(A) \to H^{i+1}(B) \to \cdots,$$

Remark that the composition

$$A \to B \to C(f)$$

is not zero, but is nullhomotopic.

Now comes a key lemma (Proposition 2.16 in Huybrechts). First one writes down a funny morphism

$$g : A[1] \to C(\tau)$$

as follows. Note that

$$C(\tau)^i = B^{i+1} \oplus C(f)^i = B^{i+1} \oplus A^{i+1} \oplus B^i.$$

and $A[1]^i = A^{i+1}$, so the morphism will be

$$g = (-f^{i+1}, 1, 0).$$

The lemma states that this morphism is an isomorphism in the homotopy category $K(\mathcal{A})$. The morphism $C(f) \to A[1]$ completing g to a homotopy equivalence is the projection to the middle factor. Actually the lemma says more, it says that g fits into a homotopy commutative diagram

$$
\begin{array}{ccccccc}
B & \longrightarrow & C(f) & \longrightarrow & A[1] & \longrightarrow & B[1] \\
\| & & \| & & \downarrow{\scriptstyle g} & & \| \\
B & \longrightarrow & C(f) & \longrightarrow & C(\tau) & \longrightarrow & B[1]
\end{array}
$$

Proposition 33.1 (Prop 2.17). *Given*

$$
\begin{array}{ccc}
C & & A \\
& \searrow{\scriptstyle g} \quad \swarrow{\scriptstyle q} & \\
& B &
\end{array}
$$

with q qis, there exists

$$
\begin{array}{ccc}
& C_0 & \\
{\scriptstyle qis}\swarrow & & \searrow{\scriptstyle \cdot} \\
C & & A \\
\searrow{\scriptstyle g} & & \swarrow{\scriptstyle q} \\
& B &
\end{array}
$$

which is commutative in the homotopy category $K(\mathcal{A})$.

Proof. We have the triangle (What's a triangle?)

$$A \to B \to C(q) \to A[1]$$

and we also form the cone on $\tau \circ g : C \to C(q)$, wich gives the triangle

$$C \to C(q) \to C(\tau \circ g) \to C[1].$$

We rotate the triangle to get

$$C(\tau \circ g)[-1] \to C \to C(q) \to C(\tau \circ g),$$

and match up with the first one:

$$
\begin{array}{ccccccc}
C(\tau \circ g)[-1] & \longrightarrow & C & \longrightarrow & C(q) & \longrightarrow & C(\tau \circ g) \\
\Big\downarrow ? & & \Big\downarrow g & & \Big\| & & \Big\downarrow ? \\
A & \longrightarrow & B & \longrightarrow & C(q) & \longrightarrow & A[1]
\end{array}
$$

It is starting to look like our desired object C_0 is $C(\tau \circ g)[-1]$, we just need to find the map (?) and prove it is qis.

But we know there is a natural map $C(\tau \circ g) \to C(\tau)$, and we know from the key lemma that $C(\tau)$ and $A[1]$ are isomorphic in the homotopy category. So this gives us our map. Why is it a qis? Since q is a qis, the cohomology groups of $C(q)$ all vanish, this is because of the long exact sequence associated with the bottom line. But then looking at the top line, we see that $C(\tau \circ g)[-1] \to C$ must be a qis. So we are done. $\qquad\square$

Thus the construction of the composition of two roofs, only makes sense in the homotopy category $K(\mathcal{A})$. It is now time to describe the equivalence relation on roofs. Two roofs

are equivalent if there exists

in which the squares commute up to homotopy, and (Not sure which) the morphism $C_0 \to C_1$ is qis or the composition $C_0 \to C_1 \to A$ is qis.

Maybe we could check that quasi-isomorphisms are invertible in this category now.

34 Derived functors

Comment that the main purpose of derived categories is to provide a convenient formalism for derived functors. That is, given an additive functor $F : \mathcal{A} \to \mathcal{A}'$ between abelian categories, we would like to construct an additive functor

$$RF : D(\mathcal{A}) \to D(\mathcal{A}')$$

with the property that the cohomology objects $H^i(DF(A))$ recover the derived functors $RF^i(A)$. Since a functor always sends isomorphisms to isomorphisms,

It's pretty clear that

The solution to this problem is actually pretty lame. One essentially declares that $RF(A)$ is the complex

$$RF(A)^i = F(A^i),$$

only for objects A that lie in a certain small part of the category $D(\mathcal{A})$.

34.1 Some tricky points

Let F be a left exact functor, and RF its right derived functor. If you want to apply RF to a complex A concentrated in degree 0, you resolve it by an injective complex $A \to I$ and hit that with the functor F object-by-object. What about applying RF to a true complex A (bounded below, say)? Since the complex is bounded below, one can map the first nonzero term to an injective object, and inductively build a complex $I^\bullet$ of injectives such that $A \to I$ is an injective resolution. Then $RF(A)$ is defined by applying F object-by-object to I.

There is a spectral sequence for this. On the first page we have $E_1^{p,q} = RF^q(A^p)$, with horizontal maps $RF^q(A^p) \to RF^q(A^{p+1})$ the ones induced from the complex $A^\bullet$.

So far so good, but how does one *construct* this spectral sequence? It comes from a double complex called the Cartan-Eilenberg complex. You might think the construction is as follows: Grab an injective resolution $I_{(p)}^\bullet$ of each A^p, and put them vertically on top of each A^p, i.e., $I_{(p)}^q$ in position (p, q). Or rather maybe choose these resolutions inductively, starting from the lowest nonzero degree of A, and choosing each with maps from the previous. Whatever. This is NOT the definition of Cartan-Eilenberg resolution given in Weibel [**?**, Definition 5.7.1]. Instead one chooses resolutions of the boundary subgroups $B^p \subset A^p$ and of the cohomology groups $H^p = Z^p/B^p$, and does some horshoe lemma trick to stitch them together. The resulting columns $I^{p,\bullet}$ ARE injective resolutions of the A^p, but it appears you are not allowed to simply choose any such set of resolutions you wish. Tricky.

Some discussion on Mathoverflow `https://mathoverflow.net/questions/89598/higher-cartan-eilen` led to the following puzzle. The book of Benson [**?**, Definition 2.7.4] defines projective resolution as follows:

Suppose that C is a chain complex of left A-modules, bounded below. Then a projective resolution of C is a chain complex P of projective left A-modules, bounded below, together with a map of chain complexes $P \to C$ which is an isomorphism on homology.

According to the Mathoverflow discussion, there is an erratum to this book in which Benson concedes that "The definition is incorrect; the Cartan-Eilenberg theory of proper resolutions cannot be avoided here.", and the problem is that complexes of projectives are not necessarily projectively cofibrant.

On the other hand [**?**, Theorem 10.5.6] seems to assert that for bounded below/above complexes it's OK to grab a quiso complex of injective/projective objects. So I don't see what's wrong with the definition as stated by Benson.

Question: What's the deal with Benson's erratum? The discussion around it seems to suggest that one is not supposed to use arbitrary injective resolutions to compute derived functors, only resolutions of the form $Tot(I^{\bullet,\bullet})$. I can see this being true for unbounded complexes, but for lower-bounded ones I think this difficulty does not arise.

In Spaltenstein's paper [**?**, p. 124] it is explained that an unbounded-in-both-directions acyclic complex of projectives might not be homotopic to the zero complex. If we consider both of them resolutions of the zero complex then we are in trouble. He introduces the notions of K-injective and K-projective complexes to substitute the bad notion of "complex all of whose terms are projective/injective". Again, I see this as being a problem about complexes unbounded-in-both-directions.

35   Perverse sheaves

Notion of stratification of an algebraic variety. The Whitney properties.

We fix a stratification $X = \cup S_i$ of a variety X. A sheaf $\mathcal{F}$ on X is said to be constuctible (relative to the chosen stratification) if for each i the restriction $\mathcal{F}|_{s_i}$ is a locally constant sheaf with finite dimensional vector spaces for stalks. Recall $D^b(X)$ is a derived category of bounded complexes of arbitrary sheaves on X. The constructible derived category $D^b_c(X)$ is the full subcategory of $D^b(X)$ consisting of objects $\mathcal{F}^\bullet$ whose cohomology sheaves (not Cech cohomology, here we just mean kernel over image in the literal sense) are all constructible sheaves.

In theory this category contains complexes with terms very far from being local systems, since only the cohomologies are required to be constructible. Indeed it seems to me (find explicit reference) that a primordial example of this is the de Rham complex resolving the constant sheaf of a smooth variety

$$qis : \wedge^\bullet \Omega_X \to \underline{\mathbb{C}}_X.$$

On the other hand see the discussion in [**?**, Section 4.1]: the constructible sheaves on X form an abelian subcategory $Sh_{constr}(X)$ of the category $Sh(X)$ of all sheaves on X. One can then form the derived category $D^b(Sh_{constr.}(X))$ and it turns out there is a naatural functor

$$D^b(Sh_{constr.}(X)) \to D_c^b(X).$$

Apparently [**?**, Theorem 4.1.4] if X is a complex algebraic variety, and the stratification is by Zariski subsets, then this functor is an equivalence. This would not be the case for a general analytic stratification.

The slides [**?**] are quite advanced, but seem to be good.

Consider a variety X with an algebraic Whitney stratification (such as $X = G/B$ with the stratification by Bruhat cells). Let $U \subset X$ denote the open cell and $i : U \to X$ the inclusion. The idea is that there is a "preferred" way to extend sheaves from U to X, different from the two pushforwards i_* and $i_!$. Two questions and an observation arise:

- (Observation) The resulting objects are not sheaves but rather objects of the derived category $D_c^b(X)$ of constructible sheaves on X

- (Question 1) What actually is the "extension"? It is called intermediate extension and its symbol is $i_{!*}$. Definition below.

- (Question 2) What is preferred about this extension?

The image of $i_{!*}$ lands in a subcategory of $D_c^b(X)$, and the restriction of $i_{!*}$ to local systems on U lands in a sub-subcategory which are called perverse sheaves on X, this is an equivalence.

36   Mixed Hodge modules

Not sure how much we will get into this.

But the volume on motives (Jannsen, Kleiman) has a nice little chapter on the definition of Hodge structure, etc.

NOT SURE WHERE TO PUT THIS SECTION

37   Symplectic structure of projective varieties

Definition 37.1. Let V be a vector space over the field of complex numbers $\mathbb{C}$. A Hermitian form on V is a function

$$(\cdot, \cdot) : V \times V \to \mathbb{C}$$

which is

- $\mathbb{C}$-linear in the second argument, i.e., $(u, a_1 v_1 + a_2 v_2) = a_1(u, v_1) + a_2(u, v_2)$,

- conjugate symmetric, i.e., $(v, u) = \overline{(u, v)}$.

The standard Hermitian form on $\mathbb{C}^n$ is given by

$$((u_1, \ldots, u_n), (v_1, \ldots, v_n)) = \sum_i \overline{u}_i v_i.$$

A nifty thing about the standard Hermitian form is that $(v, v) \in \mathbb{R}$ and $(v, v) \geq 0$ with equality only for $v = 0$.

Definition 37.2. Let W be a vector space over the field of real numbers $\mathbb{R}$. A symplectic form on W is a function

$$\omega(\cdot, \cdot) : W \times W \to \mathbb{R}$$

which is

- $\mathbb{R}$-linear in both arguments,

- skew-symmetric, i.e., $\omega(v, u) = -\omega(u, v)$,

- Non-degenerate, i.e., for each $u \in W$ the set $\{v \in W \mid \omega(u, v) = 0\}$ is not all of W.

Let V be a complex vector space. It becomes a real vector space (of twice the dimension) by restricting scalars from $\mathbb{C}$ to $\mathbb{R}$. Denote this space W. If () is a *non degenerate* Hermitian form on V then $\omega(u, v) = \Im(u, v)$ is a symplectic form on W.

So $\mathbb{C}^n$ acquires a standard symplectic form from its standard Hermitian form.

Now let V be a complex vector space with a nondegenerate Hermitian form. Let's think about the complex projective space $\mathbb{P}V$. The points of projective space are lines through the origin. Each vector $v \neq 0$ determines such a line uniquely, which we denote $[v]$.

Exercise 37.1. *Show that $\mathbb{P}^N$ is a compact topological space. You might enjoy proving this by doing the quotient in two stages:*

$$\mathbb{C}^{N+1} \backslash \{0\} \to (\mathbb{C}^{N+1} \backslash \{0\}) / \mathbb{R} \to (\mathbb{C}^{N+1} \backslash \{0\}) / \mathbb{C} = \mathbb{P}^N.$$

Definition 37.3. A symplectic manifold is a smooth real manifold M together with a smoothly varying symplectic form ω_p on the tangent space $T_p M$. Also the form should be closed.

What does smoothly varying symplectic form mean, concretely? Well, on any coordinate chart of an atlas of M, with coordinates $x_1, \ldots, x_n$, we have the vector fields $\partial/\partial x_i$, which form a basis of $T_p M$ at each p. Relative to this basis, ω_p has a Gram matrix. The entries of the Gram matrix are then functions of p. We are just requiring that these functions are all smooth.

Closed means closed as a differential 2-form.

It turns out $\mathbb{P}V$ is a symplectic manifold. Let $[v] \in \mathbb{P}V$ and let ξ_1 and ξ_2 be two tangent vectors at $[v]$. Thinking of tangent vectors at $[v]$ as equivalence classes of smooth curves, we can lift representatives of ξ and ξ_2 to smooth curves $\gamma_1(t)$ and $\gamma_2(t)$ passing through v. Convince yourself that if $w = \alpha v$ is another representative of $[v]$, then $\alpha \gamma_i(t)$ is a representative of ξ_i through w. If we add the proviso that our $\widetilde{\xi}_i = \gamma_i'(t)$ be *orthogonal* to v with respect the Hermitian form on V, then

$$\omega_{[v]}(\xi_1, \xi_2) = \frac{(\widetilde{\xi}_1, \widetilde{\xi}_2)}{(v, v)}.$$

is well-defined, and it is a symplectic form. See [**?**].

Now let X be a smooth projective algebraic variety, with embedding $X \to \mathbb{P}V$ into projective space. As a general rule if $N \subset M$ is a submanifold, and M is symplectic, the restriction of ω_p to $T_p N$ might be degenerate, so N might not be symplectic.

Exercise 37.2. *Use the magic of Hermitian property to show that $X \subset \mathbb{P}V$ is symplectic.*

Now suppose our smooth projective variety X carries an action of the complex Lie group G.

Moment map